Caractérisation structurale, enzymatique et biophysique d'un complexe peptidase piezo-thermophile issue de l'archaea marine abyssale Pyrococcus horikoshii

Eva Rosenbaum

Université Joseph Fourier - Grenoble I

Ecole Doctorale de Physique

Thèse

Présentée et Soutenue Publiquement le 3 Décembre 2008 par

Eva Rosenbaum

pour l'obtention du grade de

Docteur de l'Université Joseph Fourier - Grenoble I

Spécialité : Biophysique

Structural, biophysical and enzymatic characterization of a piezo-thermophilic peptidase complex from the deep sea archaeon *Pyrococcus horikoshii*

Caractérisation structurale, enzymatique et biophysique d'un complexe peptidase piezo-thermophile issue de l'archaea marine abyssale *Pyrococcus horikoshii*

Membres du Jury

Dr. Giuseppe Zaccai	Président
Dr. Catherine Royer	Rapporteur
Dr. Marco Moracci	Rapporteur
Dr. Thomas Gutberlet	Examinateur
Dr. Bruno Franzetti	Directeur de Thèse
Dr. Frédéric Vellieux	Directeur de Thèse

Thèse préparée à l'Institut de Biologie Structurale J.-P. Ebel, CEA/CNRS/UJF, Grenoble

Caractérisation structurale, enzymatique et biophysique d'un complexe peptidase piezo-thermophile issue de l'archaea marine abyssale *Pyrococcus horikoshii*

Récemment Franzetti et al. ont découvert un nouveau type de protéases autocompartimentées indépendantes d'énergie dans les archaeas [1]. Les particules ont été appelées TET pour leurs structures tridimensionnelles tétraédriques. Les TETs forment de grands complexes d'un poids moléculaire d'environ 500kDa. Leur rôle dans l'organisme est pourtant inconnu. Dans *P. horikoshii*, une Archaea hyperthermophile de la mer profonde, trois protéases TET ont été identifiées (PhTET1, 2 et 3). Nous avons exprimé et purifié PhTET3 recombinante. L'enzyme a été caractérisée biochimiquement et nous avons déterminé la structure d'un complexe de PhTET3 de 12 sous-unités par cristallographie aux rayons X. Afin de mieux comprendre son rôle physiologique potentiel et de s'assurer pourquoi il y a trois protéases TET dans *P. horikoshii*, la structure et les propriétés enzymatiques de PhTET3 ont été comparées à celles de deux autres protéases TET déjà caractérisées. Puisque l'autocompartimentage joue un rôle important dans le fonctionnement et la régulation des protéases, les facteurs commandant l'oligomérisation de PhTET3 *in vitro* ont été étudiés par ultracentrifugation analytique et diffusion de neutrons aux petits angles. Finalement, dans des états physiologiques de mer profonde, l'enzyme est exposée à la haute température (jusqu'à 100°C) et à la haute pression. Afin d'étudier les limites de la stabilité de grands assemblages macromoléculaires, la structure à basse résolution et l'activité enzymatique de PhTET3 ont été mesurées à hautes pressions et à hautes températures en utilisant la diffusion des rayons X aux petits angles et la spectrophotométrie à haute pression. Au total, ces études ont indiqué que les protéases TET de *P. horikoshii* forment un système intégré de dégradation de peptides et que PhTET3 montre une stabilité exceptionnelle à haute pression et à haute température aussi bien que des propriétés enzymatiques associées aux conditions environnementales.

Mots clés : protéase TET, extremophiles, stabilité des protéines, complexes macromoléculaires, *Pyrococcus horikoshii*, pression hydrostatique, diffusion aux petits angles, ultra centrifugation analytique, cristallographie aux rayons x, spectrophotométrie sous haute pression

Structural, biophysical and enzymatic characterization of a piezo-thermophilic peptidase complex from the deep sea archaeon *Pyrococcus horikoshii*

Recently Franzetti et al. discovered a new type of energy-independent self-compartmentalized proteases in Archaea [1]. The particles were named TET for their tetrahedral three-dimensional structure. The TETs self assemble as large molecular weight complexes of about 500kDa. Their role in the organism is yet unknown. In *P. horikoshii*, a hyperthermophilic Archaea from the deep sea, three TET proteases have been identified (PhTET1, 2 and 3). We have expressed and purified recombinant PhTET3. The enzyme was characterized biochemically and we determined the structure of a PhTET3 12 subunit complex by x-ray crystallography. In order to gain insight into its potential physiological role and to ascertain why there are three TET proteases in *P. horikoshii*, the structure and the enzymatic properties of PhTET3 have been compared to the two other TET proteases that were characterized before. As self-compartmentalization plays an important role in functioning and regulation of proteases, the factors controlling PhTET3 oligomerization *in vitro* have been studied by analytical ultracentrifugation and small angle neutron scattering. Finally, under physiological deep sea conditions, the enzyme is exposed to high temperature (up to 100°C) and high pressure. In order to study the limits of stability of large molecular weight assemblies, the low-resolution structure and the enzymatic activity of PhTET3 have been monitored under high pressures and temperatures using small angle x-ray scattering and high-pressure spectrophotometry. Taken together, these studies revealed that the TET proteases of *P.horikoshii* form an integrated peptide degradation systems and that PhTET3 exhibits unusual stability under high pressure and temperature as well as environment-associated enzymatic properties.

Key words: TET protease, extremophiles, protein stability, macromolecular complexes, *Pyrococcus horikoshii*, hydrostatic pressure, small angle scattering, analytical ultra centrifugation, x-ray crystallography, high-pressure spectrophotometry

Acknowledgements

First of all I would like to thank the members of the jury for accepting to evaluate the manuscript, for participating in the thesis defense and for their helpful comments and corrections. Thank you Catherine Royer, Marco Moracci, Joe Zaccai, Thomas Gutberlet, Bruno Franzetti and Frédéric Vellieux.

Without my two thesis supervisors Bruno Franzetti and Fred Vellieux none of this would have been possible. So thank you Bruno and Fred for all your encouragement, your support and your enthusiasm for this project.

I am also very grateful to Asun Durá for all the input and support, especially concerning the biochemistry parts of the project.

A big thank you to Mylène Ferruit and to Frank Gabel for helping me in so many different ways, for your dedication to this project and for many helpful discussions.

I am indebted to all the people that shared their expertise in different biophysical techniques with me. Thank you Christine Ebel, Aline Appourchaux, Cécile Cléry-Barraud, Patrick Masson, Stéphane Marchal, Reinhard Lange, Stéphanie Finet, Eric Girard, Richard Kahn and Joe Zaccai for all your help and support.

Thank you also to the teams of ID02, D22, ID23 and ID27 at the ESRF and the ILL. It has been a pleasure to use these instruments.

I would also like to thank the members, past and present, of the LBM for all the good times we had inside and outside the lab. Thank you Benoît, Nico, Marion, Katy, Mylène, Aline, Amédé, Jaques-Philippe, Renata, Iza, Antosia, Asun, Manu, Dorian, Vincent, Arnaud, Marcus, Domi, Frank, Judith, Bruno, Martin, Fred and Christine.

Je suis aussi extrêmement reconnaissante envers mon équipe de basket et tous les autres gens de l'ESSM. Merci pour tout.

Most of all I would like to thank my parents and Marcus. Thank you for your support, your faith in me and for helping me make the right choices. You're awesome!!!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Archaea

1.1.1 The tree of life

In 1990 Woese and coworkers [2] proposed the introduction of a phylogenetic tree that comprised as its highest hierarchical level three domains of life: Bacteria, Archaea and Eucarya (Figure 1.1). The relationship between the different domains was inferred by analysis of ribosomal RNA sequences [2]. The domain of Archaea was first reported by Woese and Fox in 1977 [3]. Bacteria and Archaea are both prokaryotes, but are sufficiently distinct to be considered as two different domains of life, even though their metabolic genes seem to be related. In others aspects, such as for proteins involved in translation and transcription, archaea are more similar to eukaryotes than they are to bacteria [4, 5]. At first it was thought that this new domain of life comprised only organisms adapted to conditions that could have been found on earth three or four billion years ago. As more species of archaea were discovered it seemed as if they were only present in environments that could be classified as extreme from the human point of view, such as high or low temperature, high salinity or lack of light and oxygen [6]. However, new developments in metagenomics offer the possibility to study the world of uncultivable microorganisms. Recent studies show that archaea are abundant in most ecosystems on earth, even in "non-extreme" environments [7, 8]. In addition the organisms adapted to the most extreme physical and chemical conditions on earth are found among the archaea [9].

Today the division into the three domains is generally accepted. However whether there is a common root for all three domains of life and the nature of this last universal common ancestor (LUCA), if there is one, is still under debate ([10, 11] and references therein). To gain insight into the conditions

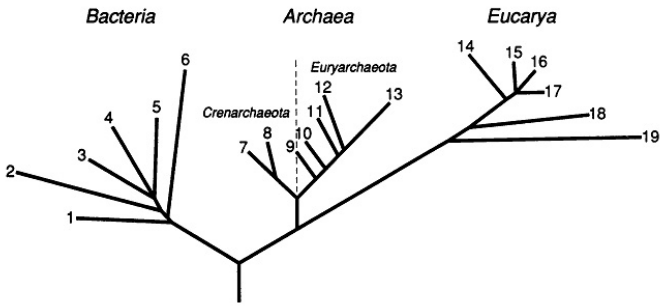


Figure 1.1: The universal phylogenetic tree as proposed by Woese et al. [2]. The three domains of life branch off from a common ancestor that Woese et al. located between the archaea and the bacteria.

under which life on earth evolved, several approaches have been used. Geology gives insight into the conditions that could have reigned on earth at the onset of life about 4 Ga[1] ago [12, 13]. Until about the same time the earth was subject to heavy meteriote bombardment [12, 13, 14]. Individual impacts would have heated the oceans on earth to over 100 °C [12]. At the same time the sun was only about 75% as bright as it is today, but emitted large amounts of radiation [13]. Therefore the surface of the earth could have been covered by a more or less thick layer of ice, locally broken by hotspots created by volcanism, and exposed to high levels of radiation [10, 13]. Alternatively the presence of greenhouse gases in sufficient amounts could have heated the surface of the earth, leading to water temperatures of up to 100 °C in the ocean for millions of years [12]. These harsh conditions are favorable for the reactions that probably led to the formation of the first biological molecules, but they are very unfavorable to the stability of these molecules [15]. It has therefore been suggested that life should have evolved in the depth of the ocean where the water layer could have shielded the first organisms from radiation and where high pressure and high temperature close to hydrothermal vents prevail [13]. If the most ancient forms of life were adapted to high temperature, they could also have survived the occasional boiling of the ocean due to asteroid impacts [12]. However it has been suggested that the earliest organisms used RNA to store genetic information and to catalyze reactions [16]. In contrast to DNA, RNA is much less thermostable and it is assumed that the upper temperature limit for correct folding of catalytic RNA molecules is 50 °C [17, 18]. This observation seems to contradict the theory of a hot origin of life. However Tehei et al. [19] were able to show that

[1] 10^9 years

high salt concentrations can stabilize tRNA molecules at high temperature. Together with research data suggesting that early oceans were more saline [20], this could present a solution of how RNA could have been used by early hyperthermophiles [19]. Furthermore it has been shown that under laboratory conditions mimicking those in hydrothermal vents, amino acids are generated abiotically, supporting the point of view that at least part of the evolution of early forms of life took place under these conditions [15]. However the conditions under which life emerged first do not necessarily mean that LUCA, the LAST Universal Common Ancestor, lived under the same conditions [21, 22]. Therefore other methods to gain information on LUCA have to be used. One of the first approaches to determine its nature was the study of phylogenetic trees. Phylogenetic trees based on the analysis of rRNA such as those presented by Stetter [14] and Pace [23] show a cluster of hyperthermophiles around the root of the tree, suggesting that the organism at the root of the tree, LUCA, was a hyperthermophile as well [14, 17]. On the other hand phylogenetic trees that are constructed using metabolic genes instead of genes from the nucleic acid based information processing system often do not show the same organization [23, 24]. Moreover lateral gene transfer between and within domains makes the construction of valid phylogenetic trees even more difficult [23].

Another approach to the question about the nature of the earliest life forms on earth is the analysis of the amino acid sequences of proteins that are highly conserved throughout all forms of today's life. DiGiulio devised a thermophily [25] and a barophily [26] index for the 20 amino acids, which indicate if an amino acid is preferentially used in (hyper-)thermophilic or barophilic organisms. DiGiulio was able to show that there is a correlation between the number of codons attributed in the genetic code to a given amino acid and its thermophily or barophily rank [25, 26]. Assuming that the number of codons in the genetic code attributed to each amino acid reflects the frequency at which the amino acid was used by the organisms living at the time the genetic code originated, these organisms should have been baro- and thermophiles [25, 26]. Furthermore using sets of paralogous [22] and orthologous [27] proteins, ancestral protein sequences for several proteins have been calculated. Analyses of the thermophily index of these ancestral sequences indicates that the LUCA was a hyperthermophile [22, 27]. What is more these analyses are independent of the hypothetical location of the root on the tree of life [22]. So far no generally accepted theory for the temperature at which LUCA lived has been developed. To gain further insights into a possible origin of life at high temperatures and at high pressure, it is important to study macromolecules from extant hyperthermophilic and barophilic organisms in order to explore the limits of their stability.

1.1.2 Pyrococcus horikoshii

The archaea *P. horikoshii* was first described by González et al. [28]. It was isolated from the Okinawa Trough in the Pacific Ocean at a depth of 1395m. The authors reported an optimal growth temperature of 98 °C, but growth was observed from 80 °C to 102 °C. In [29] K. Horikoshi reported an optimal growth temperature of 95 °C at pressure conditions between 0.1 and 15MPa[2]. At higher pressure (30MPa) the optimal growth temperature was 100 °C. Horikoshi also reported that at higher growth temperatures (100 °C and 103 °C) *P. horikoshii* showed barophilic characteristics whereas at lower growth temperatures (90 °C and 95 °C) a barotolerant behavior was observed.

1.2 Adaptation to extreme conditions

1.2.1 Extremophiles

As their name implies extremophiles are organisms that optimally grow under extreme conditions. These conditions can include low or high temperature, high salinity, extremes of pH and hydrostatic pressure. It should however be noted that the term "extreme condition" is relative and depends on the point of view. About 70% of the earth's surface is covered by oceans, where habitats of low or high temperature (e.g. close to hydrothermal vents), high hydrostatic pressure and high salinity are abundant [30, 31]. To be able to survive in these conditions organisms have developed several strategies of adaptation [30]. The most obvious approach is to avoid the destabilizing factor by maintaining moderate conditions inside the cytoplasm or by evacuating extreme chemical factors from the cytoplasm by transmembrane pumps [32, 33]. It is however difficult to apply these strategies of avoidance to high temperature or high pressure for example. Therefore the macromolecules of most extremophiles have to resist the conditions in which the organisms live [30]. For some conditions it has been possible through extensive studies to gain some insights into the mechanisms of adaptation on a macromolecular level, at least concerning the stabilization of monomeric proteins [34, 35, 36]. However the mechanisms underlying the stability of proteins under extreme conditions are still poorly understood and even more so those that stabilize large oligomeric complexes [37].

[2] or 1 to 150 bar, corresponding to a depth of 0 to 1500m in water

1.2.2 Thermophilic Adaptation

P. horikoshii (see section 1.1.2) displays optimal growth at temperatures close to the boiling point of water. Therefore all biological macromolecules of the organism have to be adapted to function at these temperatures. The ability of a protein to fulfill its biological function is closely linked to its three-dimensional structure and to its dynamics [38]. The activity and the structure of enzymes from species living at different temperatures have been thoroughly studied and a number of mechanisms of thermal stabilization have been identified.

The structure of a protein is defined by a range of interactions between different parts of the protein and between the protein and its solvent [39]. It has become evident that the differences in thermal stability and optimal catalytic temperature between orthologous proteins are due to only small changes in the sequence and in the three-dimensional structure [40, 41, 33, 42]. These modifications include changes in the size of the hydrophobic core of the protein, the number of hydrogen bonds and ion pairs, conformational strain, intersubunit interactions and in the size of loops and of the terminal regions in the protein [39, 43, 41, 33, 44]. However it has been shown by comparing groups of orthologues that the observed increase or decrease in thermostability is the sum of a number of different modifications [40, 43, 41, 45]. The modifications listed above are trends observed on a large number of datasets and changes opposite to these rules may be observed depending on the area of the protein where the change takes place ([46, 44] and references therein). By comparing the x-ray crystallographic structure of an extremely thermostable malate dehydrogenase (MalDH) with those of less thermostable homologs, Irimia et al. [47] were able to identify several of the stabilization mechanisms mentioned above in the enzyme. The hyperthermostable MalDH displays a larger number of ion pairs and hydrogen bonds, shorter external loops, a smaller total volume and a reduced accessible surface area than its less thermostable homologs [47].

It is however important to note that, in order to catalyze a reaction, an enzyme does not only need to keep a certain three dimensional structure that allows binding of substrate and cofactor, but that some degree of flexibility of the protein is required for the reaction to take place [44, 39, 41]. Hence adaptation of an enzyme to fulfill its role in the cell at a specific temperature is always a compromise between stability of the structure and flexibility of the parts of the protein needed for catalysis [40, 33, 48, 49, 45]. This is illustrated by the fact that many enzymes from thermophiles or hyperthermophiles display only little or no activity at mesophilic temperatures, even though their three-dimensional structure is maintained [48, 44]. Figure 1.2

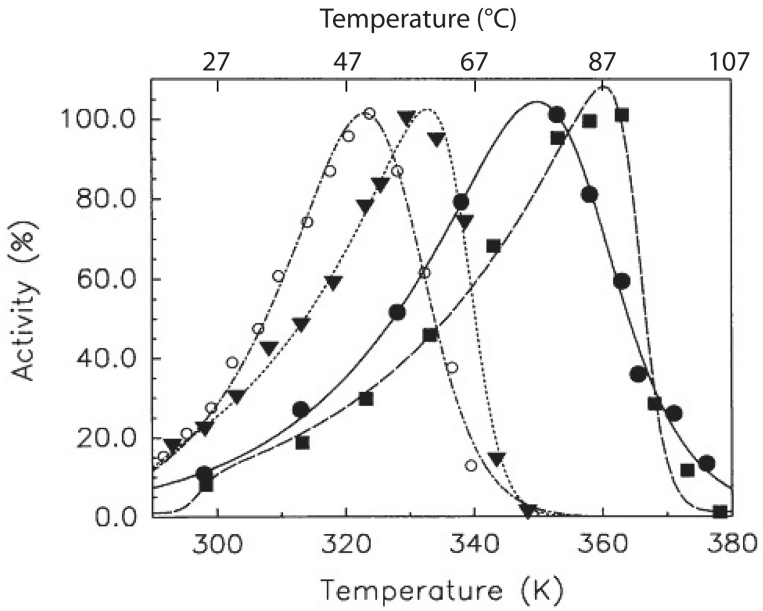


Figure 1.2: Activity at different temperatures for hydrogenases from *M. igneus* (squares), *M. jannaschii* (filled circles), *M. thermolithotrophicus* (triangles) and *M. maripaludis* from [50]. The first two organisms are hyperthermophiles, *M. thermolithotrophicus* is a thermophile and *M. maripaludis* is a mesophile. The enzymes display maximum activity at different temperatures as they are adapted to different habitat temperatures.

from [50] shows the activity of enzymes from organisms adapted to different temperatures. Whereas the enzyme from a mesophilic source is optimally active at about 50 °C the homologues from hyperthermophiles only display 30% of their maximum activity at the same temperature. Neutron scattering experiments have allowed insights into the dynamics of proteins adapted to different temperatures. The mean square atomic fluctuation and the mean force constant measured for a hyperthermophile and a mesophile show that both enzymes have a similar flexibility at their optimal activity temperature, but the hyperthermophilic protein displays a higher structural rigidity that is thought to be responsible for its enhanced stability [51].

1.2.3 Barophilic Adaptation

1.2.3.1 General considerations

A large part of the biosphere is exposed to high hydrostatic pressure. For example 62% of the volume of the global biosphere can be found at depth

exceeding 1000m [52, 53], meaning pressure of 100bar or more in water. Pressure effects on a system follow Le Chatelier's principle, stating that a system in equilibrium will minimize the perturbing effects by any external factor [54]. The equilibrium of a system with two different states is described by an equilibrium constant κ (which is denoted K in the static and k in the dynamic case). The equilibrium constant is linked to the free energy difference ΔG between the states as described in equation 1.1.

$$\kappa = e^{-\frac{\Delta G}{RT}} \tag{1.1}$$

In this R is the gas constant and T is the temperature in Kelvin. At constant temperature and pressure,

$$\Delta G = \Delta U + P\Delta V - T\Delta S$$

where U is the internal energy, P is the pressure, V is the volume and S is the entropy of the system. The volume effect of pressure on the system is determined by the volume change ΔV, because

$$\left(\frac{\partial(\Delta G)}{\partial P}\right)_T = \Delta V$$

and

$$\left(\frac{\partial(ln\kappa)}{\partial P}\right)_T = -\frac{\Delta V}{RT} \tag{1.2}$$

The volume difference ΔV between two states of a protein system is the sum of three contributions: the intrinsic contributions ΔV_{intr} due to changes in interactions between chemical groups within the protein or between the protein and other solutes, a solvation contribution ΔV_{solv} due to changes in the solvation of the protein or of other solutes and due to changes in the water structure, and a conformational contribution ΔV_{conf} due to conformational changes in the protein [55]. It is expected that the volume difference between two states is dominated by the hydration changes [54].
The change of the partial molar volume V_i of the protein with pressure is expressed in terms of the isothermal compressibility β [55]:

$$\beta = -\frac{1}{V_i}\left(\frac{\partial V_i}{\partial P}\right)_T$$

The partial molar volume of a protein can be written as the sum of three contributions [56]:

$$V_i = V_{atoms} + V_{cavities} + \Delta V_{hydration}$$

V_{atoms} and $V_{cavities}$ are the volumes of the protein's atoms and its internal cavities and $\Delta V_{hydration}$ accounts for volume changes due to protein-solvent interactions. V_{atoms} is approximately constant and thus the compressibility will depend on cavities and the hydration contribution [55]:

$$\beta = -\frac{1}{V_i}\left(\frac{\partial V_{cavities}}{\partial P} + \frac{\partial \Delta V_{hydration}}{\partial P}\right)_T$$

1.2.3.2 Pressure effects on macromolecular structures

The primary structure of a protein stabilized by covalent bonds is insensitive to pressure in the physiological range and well beyond [54]. How the weak interactions stabilizing a protein's secondary, tertiary and quaternary structure are affected by hydrostatic pressure has been studied on a number of model systems and on proteins [50, 54, 57]. It is generally accepted that solvent exposed ionic interactions tend do be disrupted by pressure as the exposure of charged groups to the solvent leads to electrostriction, thereby reducing the volume of the system [31, 54, 57, 58]. The same effect can be observed in pure water, whose pH is shifted by 0.3 during a pressure change from atmospheric pressure to 1000bar [31]. In a similar way hydrophobic interactions in a protein are solvated during exposure to high pressure as water penetrates into the hydrophobic core of the protein [31, 59, 54, 60]. The penetration of water into the protein leads to a swelling of the structure into a molten globule form and ultimately to denaturation which is also accompanied by a negative volume change [13, 61, 62]. By using Pressure Perturbation Calorimetry (PPC) it has been shown that the coefficient of thermal expansion, α, of amino acid side chains becomes more negative at low temperature with increased side chain hydrophobicity [63, 64]. This behavior indicates that the solvation of hydrophobic side chains increases the amount of structure in liquid water, which leads to an increase in volume [64]. Thus the exposure of hydrophobic residues to the solvent should become more unfavorable with increasing pressure. Stacking interactions between aromatic rings are also thought to be enhanced by pressure [54]. Finally the formation or disruption of hydrogen bonds is associated with very small volume changes that can be positive or negative [31, 59, 54]. Hydrogen bonds are therefore essentially insensitive to pressure.

When looking at the difference in free energy ΔG between the unfolded and the native state of a protein, it can be expressed as [56]

$$\begin{aligned}\Delta G = \Delta G_0 - \Delta S_0(T - T_0) - \frac{\Delta C_p}{2T_0}(T - T_0)^2 + \Delta V_0(P - P_0) \\ + \frac{\Delta\beta}{2}(P - P_0)^2 + \Delta\alpha(P - P_0)(T - T_0) + \text{higher order terms}\end{aligned} \tag{1.3}$$

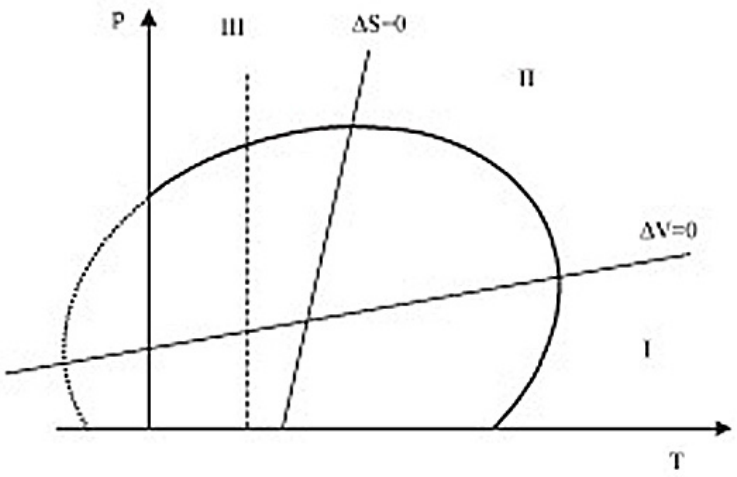


Figure 1.3: A typical pressure-temperature phase diagram for proteins. In zone I ΔV and ΔS are positive, in zone II ΔV is negative and ΔS is positive and in zone III both ΔV and ΔS are negative. Pressure unfolding at ambient temperature is usually observed in zone III [65].

In this $\Delta\beta$ is the difference in compressibility, $\Delta\alpha$ is the difference in thermal expansion and ΔC_p is the difference in heat capacity. When plotting the $\Delta G = 0$ curve in the P-T plane an elliptical shape is obtained, which reflects the significant contributions from the second order terms to ΔG [56]. In figure 1.3 a typical phase diagram for proteins is displayed (taken from [65]). In the phase diagram the area inside the curve corresponds to a state of the protein where ΔG is positive and the protein is thus in its native state. For the area outside the curve the opposite is true [65]. From the phase diagram it can be seen that in zone III the pressure at which the protein denatures increases with increasing temperature.

1.2.3.3 Experiments on protein structure under high pressure

A number of studies have been performed on monomeric and on oligomeric proteins regarding their stability under hydrostatic pressure. It has been shown that murine prion protein (mPrP) will form aggregates when exposed to hydrostatic pressure of more than 4000 bar. In another study Gorovits et al. [66] observed the dissociation of the chaperonin Cpn60 (GroEL) from *E. coli*. The 14-subunit oligomer started to dissociate at a pressure of 1300 bar. Another example of oligomer dissociation by hydrostatic pressure is the de-oligomerization of RuBisCO[3] from *R. rubrum* [67]. A small degree of disso-

[3]ribulosebisphosphate carboxylase/oxygenase

ciation was already observed at a pressure of 400 bar and at about 1200 bar about half of the dimers were dissociated. However the stability of the dimer was greatly enhanced when the inhibitor CAPD[4] was bound to the protein. In general monomeric proteins are quite resistant to hydrostatic pressure at ambient temperature and neutral pH and in many cases do not undergo denaturation until pressures of 3000 or 4000 bar [31, 13, 68], but unfolding at lower pressure has been reported [69]. Oligomeric proteins can be dissociated by much lower pressure that is within the physiologically relevant range [30, 13]. The dissociation of oligomers is often accompanied by structural changes or unfolding of the subunits [69]. A number of mechanisms explaining pressure resistance and adaptation of proteins have been proposed [70, 71]. Yet no general mechanism has been found for monomeric proteins and even less is known about the pressure stability of oligomers [72].

1.2.3.4 Effect of pressure on enzymatic activity

Pressure will not only affect the structure of proteins but also the activity of enzymes. As for the stability, effects of pressure on enzyme activity depend on the volume changes the system undergoes during the reaction. By studying the pressure dependence of the steady-state catalytic parameters K_m and k_{cat}, two volume changes ΔV_{Km} and ΔV_{kcat} can be determined. ΔV_{Km} can be associated with substrate binding so that $\Delta V_{Km} = -\Delta V_b$ and ΔV_{kcat} corresponds to the activation volume $\Delta V^{\neq}$ of the rate limiting step [55]. Negative volume changes indicate that the reaction is favored by pressure and the opposite is true for positive volume changes. In the most straight forward cases K_m and k_{cat} show a linear pressure dependence and thus the associated volume changes do not change with pressure. However a number of non-linear pressure effects on reaction rates have been observed [73]. In a multi-step reaction the rate constants of each step may have different pressure dependences, so that at a certain pressure there will be a change in the rate-limiting step, when the two profiles cross [73, 55]. Furthermore enzymatic activity can be affected by changes in the protein structure [73, 55]. The influence of compressibility changes can be expressed in terms of the activation isothermal compressibility $\Delta\beta^{\neq}$. In this case the non-linear behavior of the rate constant k with pressure can be described by a quadratic equation [55].

$$ln(k) = ln(k_0) - \frac{\Delta V^{\neq}}{RT}P + \frac{\Delta\beta^{\neq}}{2RT}P^2 \tag{1.4}$$

[4] 2-carboxy-D-arabinitol 1,5-diphosphate

Moreover pressure can indirectly influence enzymatic activity through the dissociation of oligomers and through protein denaturation [55].
It has been shown that even the activity of relatively simple, monomeric enzymes can exhibit complex pressure dependence. Whereas the activities of lysozyme and thermolysin show linear pressure dependences up to 1000 bar with the first enzyme being inhibited and the second being activated by pressure, trypsin displays a break in the activity pressure curve [73]. At first trypsin is activated up to a pressure of about 400 bar, at higher pressure the activity stays constant [73]. Octopine dehydrogenase shows an even more complex pressure dependence that has been attributed to changes in the protein structure or conformation [73]. In contrast the activity of the tetrameric carboxypeptidase from *S. solfataricus* exhibited a linear pressure dependence from 0 to 1500 bar [74].

1.2.3.5 Pressure adaptation in proteins

Since proteins from surface organisms can already resist relatively high pressures, the question whether there is a pressure adaptation in proteins arises. This question is important since it is known that most of the deep seafloor is populated by pressure-adapted microorganisms [75]. Several authors have undertaken comparative studies on homologous proteins from species living at different hydrostatic pressures to study barophilic adaptation. Nishiguchi et al [76] compared the activities of tetrameric lactate dehydrogenase (LDH) from three hagfish living at different depths and therefore at different hydrostatic pressures (see figure 1.4). The activity of all three LDH decreased with increasing hydrostatic pressure up to 1000 bar, but the inactivation was least for the species inhabiting depths of about 1000 m and strongest for the LDH from a surface organism, suggesting a baro-adaptation of the protein from the fish living at greater depths. Morita [71] reported that the volume change upon actin polymerization is smaller in actin from abyssal fish than it is in actin from surface organisms. Furthermore DNase I inhibition and the intrinsic tryptophan fluorescence of the actins from the abyssal species are less affected by pressure than that of the surface organisms. Another study on LDH from fish revealed that the protein from deep-sea species is not only more stable against pressure inactivation but also more thermostable than its surface homolog [72]. In the last two studies the authors have also attempted to explain the differences in pressure stability by the difference in amino acid sequence. Whereas Morita [71] interpreted the sequence changes between abyssal actin and other actins in reference to known actin-actin and actin-ligand interactions, Brindley et al [72] have identified a tendency in abyssal LDH to replace lysines and methionines. A similar approach has

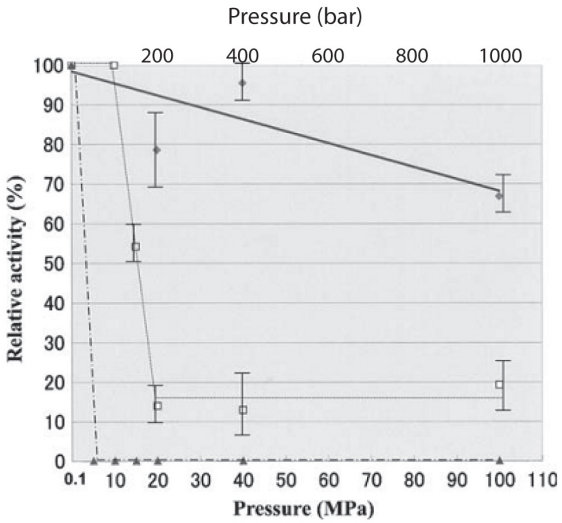


Figure 1.4: The pressure effect on the activity of three LDH from hagfish [76]. The activity of the enzyme from the deep-sea fish *Eptatretus okinoseanus* (◆) is only little affected, whereas the activity of LDH from *Paramyxine atami* (❑) that lives at depths between 250 and 400m quickly decreases with increasing pressure. The enzyme from the surface fish *Eptatretus brugeri* (▲) is inactivated by a pressure of 5MPa.

been used by DiGiulio [70], who compared sequences of proteins from two hyperthermophiles *P. furiosus* and *P. abyssi* in order to identify those amino acids that are more frequently used in the barophilic *P. abyssi*. From his findings DiGiulio calculated a pressure asymmetry index (PAI) that ranks the 20 amino acids according to their barophilicity. Of the six amino acid substitutions between the surface and the abyssal LDH that Brindley et al. [72] reported, three should increase the barostability of the protein according to the PAI of DiGiulio [70], two should have no effect and one should decrease the barophilicity. However the values reported by DiGiulio are averages over a large number of protein sequences and some substitutions against this trend might be advantageous for the baro-adaptation of the protein, especially if these changes occur in the active site, at ligand binding sites or at subunit interfaces.

An additional difficulty in studying adaptation to hydrostatic pressure is the fact that all known marine barophilic organisms are also adapted to high or low temperature[5] making it difficult to discriminate between the two adaptive mechanisms [31]. Several authors have reported an increase in the thermostability of thermophilic enzymes under high hydrostatic pressure [77, 50, 46].

[5] up to 130℃ close to hydrothermal vents, about 4℃ in the deep-sea

For example Sun et al [77] compared the effect of pressure on the thermal half-lives of hydrogenases from four methanococci: the hyperthermophile and barophile *M. jannaschii*, the hyperthermophile *M. igneus*, the thermophile *M. thermolithtrophicus* and the mesophile *M. maripaludis*. Whereas hydrostatic pressure of 500 bar led to an increase of 4.8 and 4.5 times respectively in the thermal half-lives of the two hyperthermophiles, the half-lives of the thermophile and the mesophile were lowered, but to a much larger extent in the mesophile. The authors observed the same trend for α-glucosidase from *P. furiosus* and *S. cerevisiae* and for GAPDH[6] from pig, *B. stearothermophilus* and *T. maritima*. In another study it was shown that the activity of LDH from a thermophilic bacterium was maintained up to a pressure of 2800 bar, whereas inactivation of porcine LDH began at a pressure of 1000 bar [78]. Furthermore in their study on fish LDH, Brindley et al [72] found that the LDH from the deep living fish was more thermostable, even though the organism lived at lower temperatures than the surface fish used for comparison. In view of these results the question has been raised whether baro- and thermostability in proteins are correlated [46, 79]. Temperature and pressure are thought to have antagonistic effects on protein structure with high temperature leading to enhanced flexibility of the polypeptide chain and to the loss of hydration, whereas pressure is expected to limit the motions of the polypeptide chain [55] and to increase the number of water molecules bound to the protein [54]. In fact from the elliptical shape of pressure-temperature phase diagrams for many proteins [80] it can be seen that the melting temperature of the protein is higher at slightly elevated pressure than at atmospheric pressure. Yet this should apply to mesophilic as well as to thermophilic proteins. Furthermore when comparing the temperature and pressure dependence of thermolysin with a mesophilic enzyme, vimelysin, Ikeuchi et al [81] found that the mesophilic enzyme was more stable against pressure inactivation. Thus the question whether there is a correlation between baro- and thermostability in proteins is yet unanswered and more data is needed in order to be able to gain further insight into the subject of pressure adaptation. Moreover studying the effect of pressure on protein structure can expand our knowledge on the factors that control protein stability. In this the information obtained from pressure studies is complementary to that obtained from temperature studies [82].

[6]Glyceraldehyde-3-phosphate dehydrogenase

1.3 Proteolysis

1.3.1 Importance and general mechanisms

In recent years it has been realized that the degradation of proteins in the cell is a process that is essential for the correct functioning of essential cellular processes [83]. Proteolysis is involved in processes as diverse as the cell cycle and cell division, regulation of transcription factors, antigen presentation and in diverse metabolic pathways [84]. This is achieved by tight regulation of the levels of different intracellular proteins [83]. Proteolysis is also responsible for the degradation of proteins that are misfolded [83]. This includes newly synthesized proteins that do not fold properly as well as proteins that have become unfolded or inactive [85]. External factors such as elevated temperature, high salt concentration or reactive small molecules can accelerate the denaturation of proteins [85]. Therefore proteolysis is especially important under extreme conditions that exert an environmental stress on the cell [85, 86, 87].
In all three domains of life cytosolic proteolysis is carried out by sophisticated protein degradation machines that have little in common with the small extracellular proteases that can be found in the stomach for example. In fact the presence of extracellular proteases in the cytoplasm would be lethal for any cell as proteins would be degraded quickly in an unspecific manner [85]. To avoid unspecific degradation, intracellular proteases have either very high substrate specificity or are self-compartmentalized complexes where the active sites are shielded away from the cytoplasm inside cavities within the complex [88, 89]. Access to these cavities is only possible for unfolded polypeptide chains [84, 86]. The most prominent example of such protease complexes is the proteasome system that has been found in eukaryotes, archaea and in some bacteria [90]. In eukaryotes it consists of two subcomplexes, the 20S proteasome containing the active sites and a 19S regulatory complex. The 20S proteasome consists of four stacked rings formed by seven subunits each. There are two types of subunits, α- and β-subunits. The ring at each end of the 20S proteasome is formed by α-subunits whereas β-subunits, that contain the active sites, form the inner rings [85, 86]. The regulatory complex recognizes proteins to be degraded and unfolds them so that they can access the proteolytic chamber of the 20S complex [84]. The global architecture of the 26S proteasome is conserved between archaea and eukaryotes, however the eukaryotic proteasome is much more complex and different kinds of α- and β-subunits exist [86]. In bacteria that lack the proteasome complex, other protease complexes named ClpP, HslV or Lon are responsible for the degradation of proteins into oligopeptides [90]. All the above-mentioned proteases

have in common that they are ATP-dependent. The energy is needed for substrate recognition, protein unfolding and translocation of the polypeptide chain that is to be degraded [91]. Once inside the proteolytic chamber the peptide chain is degraded processively into small peptides of 3 to 30 amino acids that are then released into the cytoplasm [92]. The size of the released peptides indicates that there must be other proteases operating downstream of the proteasome that further degrade the peptides to amino acids [92, 93]. Once proteins have been degraded into small peptides, it has been suggested that they are further cleaved by ATP-independent self-compartmentalized proteases [91]. One such protease is the Tricorn protease which has been identified and described by Tamura et al. [94]. *In vitro* the Tricorn protease together with its interacting factors F1, F2 and F3 can degrade oligopeptides released by the proteasome to free amino acids [94, 95]. However there is no data available showing that the tricorn protease fulfills the same role *in vivo*. Moreover Saric et al. [96] found that in HeLa cell extract polypeptides generated by the proteasome accumulated when the endopeptidase TOP[7] was inactive. When bestatin, an aminopeptidase inhibitor, was added peptides of 6-9 residues in length, the degradation products from TOP, accumulated. This is another indication that other peptidases act downstream of the proteasome in protein degradation [96]. Another ATP-independent self-compartmentalized protease is Gal6 / bleomycin hydrolase, which has been purified from yeast and has homologs in mammals, birds, reptiles and in bacteria [97]. Further members of the group of ATP-independent self-compartmentalized proteases, that could act downstream of the proteasome, are DppA and TPPII [98, 99]. Besides it has been shown that in cells where proteasome activity is inhibited or lost the activity of energy-independent protease complexes is highly enhanced and that some cells can survive under these conditions [99, 100, 101]. It has been suggested that the increase in activity is due to the up-regulation of the expression of these proteases [101]. This would indicate that one of the functions of the energy-independent proteases could be to compensate loss of proteasome activity [99, 100]. Although a number of proteases have been characterized biochemically, their exact role in the metabolism of the cell and their regulation are still not well understood [102]. Therefore one of the challenges in molecular and cellular biology is to study the properties of proteases *in vitro* and *in vivo* in order to further our understanding of such an essential cellular mechanism.

[7] thimet oligopeptidase

1.3.2 TET proteases

A new class of selfcompartmentalized proteases was discovered in *Haloarcula marismortui* by Franzetti et al [1]. In contrast to other known protease complexes, this new class does not form barrel-shaped assemblies, but forms dodecameric, tetrahedral complexes [103]. This new class was therefore named TET [1].

1.3.2.1 Distribution of the M42 family of proteases

In the MEROPS classification system TET proteases belong to the peptidase family M42, which is found in archaea and bacteria [104]. MEROPS is a peptidase database that classifies peptidases into families and clans according to similarity of the sequence of the catalytically active part of the protein [104]. Among the 45 archaeal genomes that have been completely sequenced, M42 peptidases are found in 35. The Tricorn protease is found only in 7 of the sequenced genomes. In three archaeal genomes neither a Tricorn protease nor a peptidase from the M42 family have been assigned and *Pyrobaculum aerophilum* is the only archaea in whose genome both a Tricorn protease and a M42 family peptidase have been detected[8]. The distribution of M42 and Tricorn proteases in a phylogenetic tree of archaea is shown in figure 1.5. It has been suggested that TET aminopeptidases could be functional homologs of the Tricorn aminopeptidase [1, 106]. This is supported by the fact that only one archaea has both a Tricorn protease and a M42 peptidase. Among bacteria M42 peptidases have been assigned in 133 of 430 completely sequenced genomes. In total only four structures of M42 peptidases have been solved, two from *P. horikoshii*(pdb IDs 1XFO, 1Y0R, 1Y0Y, 2CF4 [103, 107, 108]), one from the bacterium *Thermotoga maritima* (pdb ID 1VHO) and one from the bacterium *Bacillus subtilis* (pdb ID 1VHE). It has been shown that the two enzymes from *P. horikoshii* are TET proteases [103, 107, 108]. For the homolog from *B. subtilis* no data on the quaternary structure of the protein in solution is available, but from examining the packing in the crystal lattice it has been suggested that this enzyme exists as a dodecamer [103]. In the case of the protein from *T. maritima* a loop in the structure appears to interfere with oligomerization and therefore the protein is considered to be monomeric [103].

[8] according to a search in the MEROPS database [105] on July 24th 2008

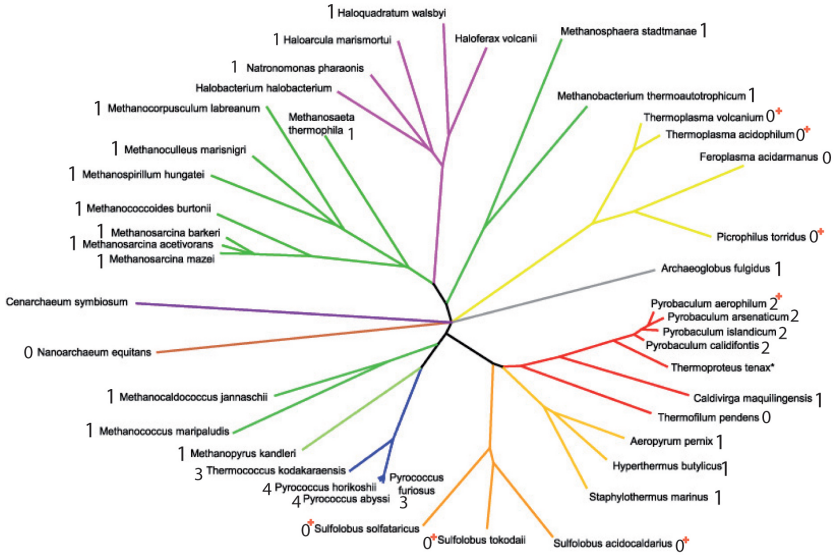


Figure 1.5: Phylogenetic tree of archaea. The number of M42 proteases identified in each genome according to the MEROPS database [105] is marked in black next to the name of the organism. Genomes in which Tricorn proteases have been identified are marked with a red cross. Note that all organisms that contain more than one M42 protease are thermophiles.

1.3.2.2 Enzymatic activity of TET peptidases

As the role of TET peptidases in the cell is yet unknown, it is important to study their biochemistry in order to gain insight into their function [106]. From the assignment of the TET peptidases into the MEROPS family M42, which belongs to the clan MH, it can be inferred that the catalytic mechanism of TET should resemble those of other proteases from family M42 or clan MH. Moreover it has been shown that the catalytic domain of PhTET1 is highly similar to AAP[9], the type example of clan MH [103, 107]. Several studies have been done on the catalytic mechanism of AAP [109, 110, 111, 112, 113]. The active sites of MEROPS clan MH peptidases contain two divalent cations, in the case of AAP zinc [114]. For the TET proteases zinc and cobalt ions have been modeled in the active sites [103, 107]. In AAP as in TET the two catalytic ions are bridged by the carboxylate group of an aspartic acid residue and by a water molecule [103, 114]. Each ion is also coordinated by the nitrogen Nϵ2 of a histidine residue. Furthermore ion 1 is coordinated by the oxygen Oϵ1 of a glutamic acid residue whereas ion 2 is coordinated by the oxygen Oδ1 of an aspartic acid residue. A catalytic glutamic acid

[9]aminopeptidase AP1

residue is present in the active site that forms a hydrogen bond with the water molecule. The active site of PhTET2 is shown in figure 1.6. According to the review by Holz [115], the first step in the hydrolysis of a peptide bond by AAP is the binding of the amide carbonyl group of the P_1 residue of the peptide to Zn1. After substrate binding the bond between Zn2 and the bridging water molecule breaks. The amide nitrogen atom of residue P'_1 forms a hydrogen bond with the catalytic glutamic acid residue and binds to Zn2. Then one of the protons of the water molecule is transferred to the catalytic glutamic acid that acts as general acid and base [112] and the Zn1-bound hydroxide ion attacks the amide carbonyl carbon of the P_1 residue. The proton from the catalytic glutamic acid residue is subsequently transferred to the amide nitrogen of the P'_1 residue, leading to the breaking of the C-N bond. It is assumed that this is the rate-limiting step in the catalysis [109, 115, 116]. Finally the products are released from the active site and a new water molecule is bound between the zinc ions [115].
Until now two TET proteases have been characterized biochemically, one from *H. marismortui* (HmTET) and PhTET2 [1, 106]. Both enzymes are strict aminopeptidases that require a free N-terminus of their substrate [1, 106]. HmTET displays a broad substrate specificity with a preference for neutral and basic residues [1]. It can process polypeptides of up to 40 amino acids, but does not cleave amino acids that are followed by a proline residue [1]. PhTET2 has a much narrower substrate specificity than HmTET and has a strong preference for small hydrophobic or uncharged residues [106]. As HmTET, PhTET2 does not act on amino acids that are followed by proline and it can act on longer peptides as well, but only up to a length of 26 amino acids [106].

1.3.2.3 Structure of the TET complex

The original shape of the TET proteases was shown by Franzetti et al [1] by electron microscopy. Russo and Baumann [103], Borissenko and Groll [108] and Schoehn et al. [107] have published x-ray crystallographic structures of homologs from *P. horikoshii*. A TET monomer consists of two domains, a catalytic domain and a dimerization domain [103, 107]. The catalytic domain is formed by a mixed eight stranded β-sheet surrounded by α-helices and a smaller four stranded anti-parallel β-sheet [107]. For dimerization the smaller β-sheet forms a number of hydrogen bonds, salt bridges and hydrophobic contacts with the dimerization domain of a second monomer [107, 108]. The dimerization domain consists of a six stranded β-sheet flanked by two short α-helices [103, 107]. Additionally in PhTET1 a disordered loop can be seen in the dimerization domain whereas in PhTET2 no electron density can be

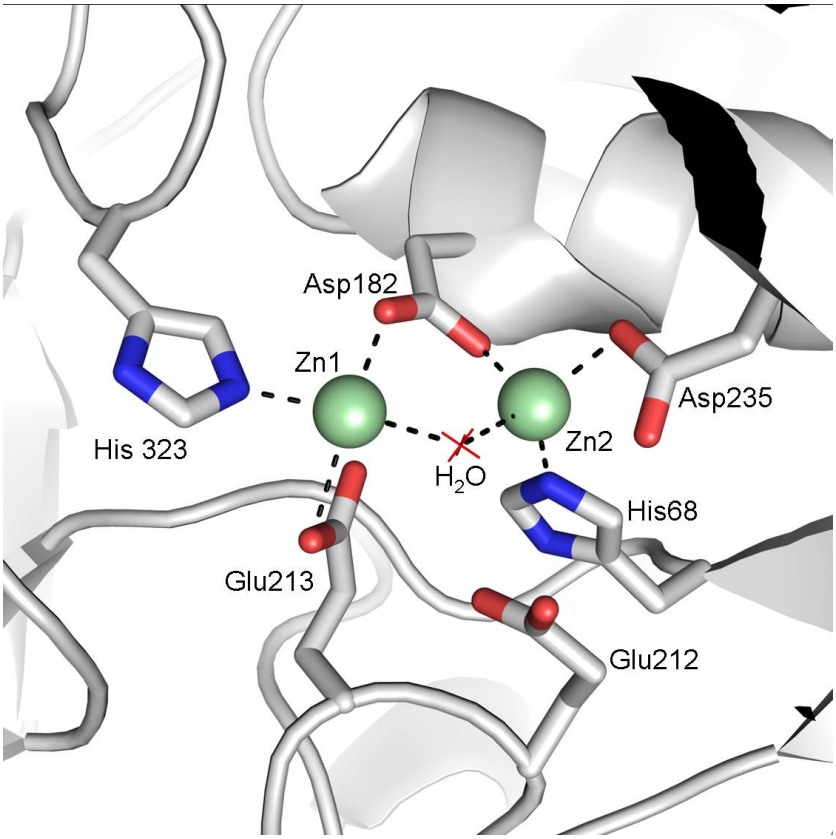
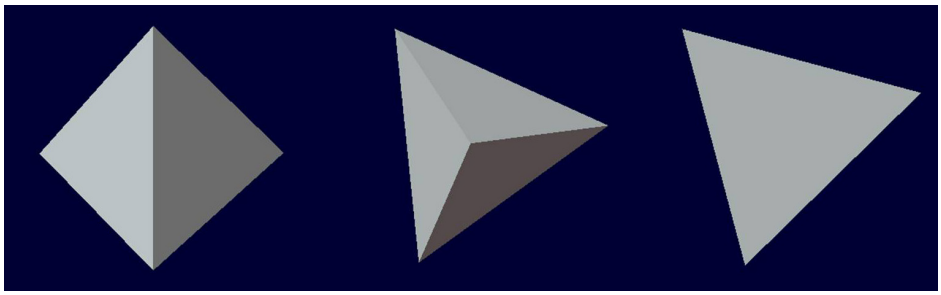
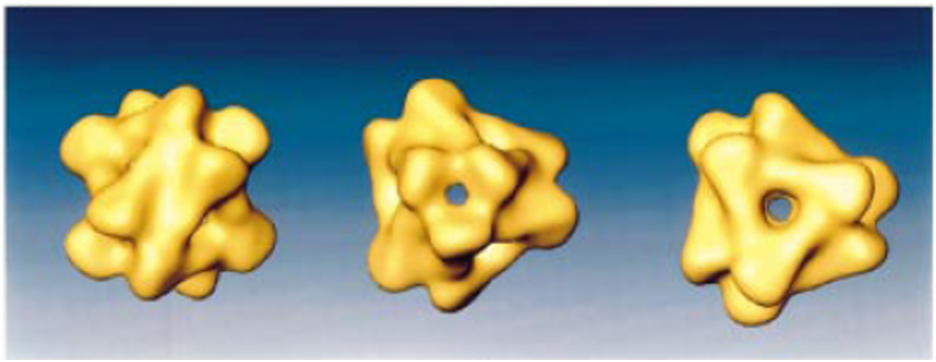


Figure 1.6: View of the active site of PhTET2 from the structure determined by Russo and Baumann [103] (pdb accession number 1XFO). The residues involved in the coordination of the catalytic ions and in the catalysis are shown as sticks. The catalytic zinc ions are shown in green and the oxygen atom of the catalytic water molecule is colored in red. In the case of PhTET2 Zn1 is coordinated by Asp 182, Glu 213 and His 323. Zn2 is also coordinated by Asp 182 as well as by His 68 and Asp 235.

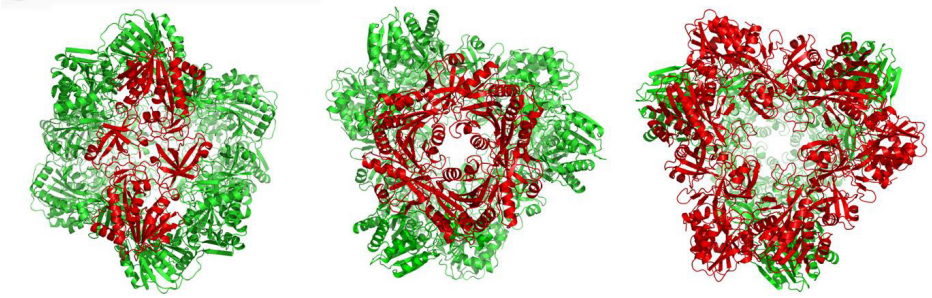
observed for the corresponding residues [107, 108]. Each monomer contains one active site with two catalytic ions that have been determined to be Zn^{2+} in PhTET2 and Co^{2+} in PhTET1 [107, 108]. The cations are tetrahedrally coordinated and the metal ligands are identical in PhTET1 and PhTET2 [103, 107].
The crystallographic models confirm that TET is a selfcompartmentalized protease. TET forms dodecameric assemblies whose building-block is probably a dimer [107]. The tetrahedral complex is formed by six dimers that build the edges of the tetrahedron as shown in Figure 1.7 [103]. The apex of the tetrahedron is formed by the catalytic domains from three subunits, forming a small opening that leads towards the interior of the complex [108]. In the case of PhTET2 this opening is almost completely blocked by Phe224 [108], but in the case of PhTET1 the opening is larger and has been proposed to serve as an exit for free amino acids generated during peptidase activity [107]. A larger opening is situated on each face of the tetrahedron, lined by six subunits [107]. These openings are thought to allow small peptides or unfolded proteins to access the interior of the protein where the active sites are located [103, 108, 107]. In the case of PhTET2 the openings lead into an internal chamber from where the twelve catalytic sites can be accessed [108]. Data obtained on PhTET1 by cryo-electron-microscopy and by x-ray crystallography revealed the structure of the interior of the TET particles in more detail and showed the existence of four channels leading towards the catalytic chambers [107]. These channels intersect in the middle of the particle, but no central cavity is identified [107]. In PhTET1 and PhTET2 the active sites are located near the apices of the tetrahedron on the same plane, close to the smaller opening [107, 108]. The charge distribution on the surface of the inside of the particle is very sophisticated and is thought to play an important role in guiding the polypeptides towards the active sites [107, 108]. For PhTET1 a second particle has been identified that assembles as an octahedron formed by 24 subunits [107]. It is proposed that this special assembly is possible in PhTET1 because of a higher flexibility of the two subunits in a dimer with respect to each other, allowing the insertion of additional dimers in the structure [107].



(a) Schematic view of the tetrahedron.



(b) EM reconstruction of the complex [1].



(c) Crystallographic model of the complex.

Figure 1.7: Organization of the TET complex. In all three lines the first view of the complex is down a two-fold axis. One of the edges of the tetrahedron is visible which is formed by a dimer. In (c) one of the dimers forming the edges is colored in red. The second view of the tetrahedron is down a three-fold axis with an apex of the tetrahedron pointing upwards. In (c) it can be seen that the apex is formed by three catalytic domains (red) arranged as a triangle. On the right the complex is again seen along a three-fold axis but in the opposite direction so that one face of the tetrahedron is visible. The face is formed by three dimers forming a triangle. The center of the triangle forms the entrance pore that gives access to the interior of the complex.

1.4 Objectives of the research work

1.4.1 Why are there several TET proteases in some organisms?

Even though it has been shown *in vitro* that TET proteases are self-compartmentalized aminopeptidases [1, 103, 106, 107], their role *in vivo* is yet unknown. Moreover in some organisms, e.g. *Haloarcula marismortui*, there is only one TET protease, whereas in other organisms, e.g. *Pyrococcus horikoshii*, several TET proteases have been identified. The structures of two TET proteases from *P. horikoshii* (PhTET1 and 2) have been solved [107, 103]. Furthermore the biochemistry of PhTET2 has been extensively characterized [106]. By studying the structure and the enzymatic activity of a third M42 protease from *P. horikoshii*, PhTET3, *in vitro* we hope to identify a potential reason for the existence of more than one TET protease in some organisms and to infer a possible role for TET proteases *in vivo*.

1.4.2 How is the oligomerization of PhTET3 controlled?

It has been shown that oligomerization and self-compartmentalization play an essential role in the functioning and the regulation of cytoplasmic proteases. Using PhTET3 as a model system, we have studied the oligomeric assembly of TET proteases. On the one hand we wanted to identify and to determine the structure of the building-block of the dodecameric assembly, which could either be a dimer or a trimer. On the other hand we have examined the resistance of the oligomer with respect to several chemical factors such as metal-content, salt-concentration and pH.

1.4.3 How does the PhTET3 oligomeric assembly behave under pseudo-physiological conditions?

P. horikoshii is a hyperthermophilic archaea isolated at a depth of 1395m. It is not known whether this depth constitutes the main habitat of the organism. However it has been shown that *P. horikoshii* can grow at high hydrostatic pressure [29]. High temperature and high pressure are known to destabilize oligomeric assemblies. Therefore it seems possible that the dodecameric form of PhTET3 is only marginally stable *in vivo*. We thus studied the low resolution structure and the enzymatic activity of PhTET3 under pseudo-physiological conditions, i.e. high temperature and high pressure.

Chapter 2

Methods Used

In this chapter I will introduce the biological methods used in this thesis to physicists and the physical methods used to biologists.

2.1 Production of pure protein solutions

Advances in recombinant DNA technology have made it possible to express genes from one organism in another organism. To achieve this, the DNA encoding the protein of interest is inserted into a vector, which in turn is transferred into the host organism which will then produce the protein. Generally an organism is chosen as a host whose molecular biology has been extensively studied. Therefore it is possible to include DNA sequences on the vector that will force the host organism to over-express the protein encoded in the vector. The host organism is then cultured and expresses the recombinant protein. Thereafter the protein of interest is separated from the proteins of the host organism. The over-expression of proteins in host organisms facilitates a lot the production of very pure protein solution, which is needed for all biophysical measurements.

2.1.1 Cloning and Transformation

As explained above, the first step in the production of recombinant protein is the insertion of the DNA coding for the protein of interest into an expression vector. Genomic DNA from another organism is cloned into a plasmid. The vector usually contains a promoter region that enhances the transcription of the gene cloned into the vector. It may also have an operator that inhibits the transcription of the cloned gene, until a small molecule is added. Furthermore it contains a gene, coding for a protein that confers resistance to an antibiotic.

Resistance to antibiotics is used later to select only those cells, that have integrated the plasmid.
Subsequently competent cells of the expression organism are transformed with the plasmid containing the gene coding for the protein to be studied. The cells used for expression can contain other additional genes, e.g. genes coding for special tRNAs. They also often lack some proteases which could degrade the recombinant protein.

2.1.2 Expression

The transformed cells then are grown until a sufficient quantity of material is obtained. If the plasmid used has some form of expression control, the molecule that induces gene transcription is added when cell growth enters the exponential phase. Cells are harvested by centrifugation before they enter the stationary growth phase. To access the protein the cells have to be disrupted. Cell membranes can be split by different methods, e.g. by using enzymes, detergents or with ultrasound. In the resulting solution the recombinant protein is mixed with the components of the host organisms. Insoluble macromolecules, such as membrane fragments, can already be removed by centrifugation at this state. If the recombinant protein is soluble, the following steps are undertaken.

2.1.3 Precipitation

In a first step it is necessary to use a simple procedure to separate the recombinant protein from the majority of the other molecules. This is mostly done by precipitation methods, e.g. salting out. Thereby the solution is treated with a chemical or by a physical means that will precipitate either the recombinant protein or macromolecules from the host organism. After centrifugation, the recombinant protein will either be found in the supernatant or in the pellet and a majority of the macromolecules from the host organism in the pellet or in the supernatant respectively, thus allowing to separate the two. In case the recombinant protein is in the pellet, it needs to be resolubilized in order to proceed with the next step.

2.1.4 Chromatography

To further separate the recombinant protein from any contaminating proteins or nucleotides from the host organism, the solution is then passed through one or more chromatography columns. The underlying idea to all chromatography is to separate the particles according to one of their properties such

as size, charge or pH. For this the solution is injected into a column that contains a solid matrix. The particles in the solution will interact more or less with the matrix of the column and will therefore be more or less retarded in their migration. The interactions with the column can also be modified by gradually changing buffer conditions in the column. The solution flowing out of the column is divided in fractions so that molecules exiting at different times will end up in different fractions.
Note that the process can be long und cumbersome, involving up to six chromatography columns and additional dialysis and concentration steps. Thus the purification of a protein, even if it is over-expressed in *E. coli*, can take more than two weeks. In order to purify proteins more rapidly it has become popular to use affinity chromatography for structural biology. For this a sequence coding for a series of amino acids that can bind to metals is added to the DNA of the protein of interest. The recombinant protein is produced with the additional amino acids attached to its N- or C-terminus. This tag can then be used to purify the protein by binding it to a chromatography column that has metal ions attached to its matrix. Even though this method allows easy purification of recombinant protein, the tag can interfere with proper protein folding or oligomerization. Additionally binding to the metal ions in the column can alter proteins that use metals as a co-factor. Hence affinity chromatography is not suitable for all proteins.

2.2 Determination of high resolution structures by X-ray crystallography

The amino acid sequence of a protein is usually known by biochemical means or by translation of the gene code. However to elucidate the secondary and tertiary structure of a protein crystallography or NMR[1] experiments are necessary.
For protein crystallography the crystal is exposed to an intense beam of hard X-rays, which is diffracted by the protein molecules. A detector records the diffracted photons and from the pattern it is attempted to infer the three-dimensional organization of the protein molecules. However, as the detector can only register the intensities of the diffraction spots, all information on the phase of the diffracted beam is lost and further information is necessary to determine the protein structure.

[1]Nuclear Magnetic Resonance

2.2.1 Protein crystallization

In theory one could do a diffraction experiment with only one protein molecule. However as x-rays interact with electrons in the atomic shell and as the main elements in biological samples (C, N, O, H) do not have many electrons, only few x-rays are scattered. Using only a single molecule one would need more than 300 years to collect enough data to determine a structure. To avoid this problem protein crystals are used. A typical protein crystal with an edge length of 100 µm contains about 10^{12} copies of the protein. A crystal can be described by defining a unit cell, containing one or more copies of the protein, that is placed on the nodes of a three dimensional lattice. This shows the highly symmetric way in which the copies of the protein are arranged in space. The quality of a crystal depends on how well a real crystal fits this theoretical description. The mosaicity of a crystal indicates how good the long range order in the crystal really is. It is also possible that the proteins forming the crystal adopt slightly different conformations that allow nevertheless the stacking of the molecules in the lattice.
To obtain protein crystals the composition of a solution containing the protein is gradually changed to reduce the solubility of the protein molecules. Small protein aggregates will form randomly, held together by weak interactions between the protein molecules. In the beginning the surface-to-volume ratio is high and only a small number of stabilizing interactions is formed. The energy associated with these interactions is not enough to compensate the diminution of entropy. Therefore the addition of a molecule to the aggregate will raise the free energy ΔG. Beyond a critical size enough interactions are formed so that the addition of further molecules will be energetically favorable. Once this critical size is passed an aggregate will become a crystallization nucleus that will continue to grow. Growth is slow in the beginning as the change in free energy when a molecule is added, $\delta\Delta G$, is still close to zero. As $\delta\Delta G$ becomes larger than kT the crystal will grow faster until the solution is not oversaturated anymore. The critical size that an aggregate has to pass to become a crystallization nucleus, ΔG_{max} and the speed of crystal growth depend on the strength of interactions between protein molecules and on the level of oversaturation of the solution. If ΔG_{max} is too high, the protein will more likely precipitate. By changing the composition of the crystallization solution the solubility of the protein and the interactions between different protein molecules can be influenced.
The most common method to achieve a slight oversaturation of a protein solution is vapor diffusion. For this method a solution of the protein at high concentration is mixed in a one-to-one ratio with a solution containing precipitants, salts, buffer and other molecules, called mother liquor. A drop

of this mixture is then placed in a closed recipient that contains an excess volume of mother liquor, in a way that the protein solution and the mother liquor are not in direct contact. As the mother liquor has a higher concentration in salt and precipitant than the protein solution, its vapor pressure will be lower and consequently water will diffuse from the protein solution towards the mother liquor, slowly raising the salt and precipitant concentration in the protein solution leading to oversaturation and consequently to crystallization. It is interesting to note that in the case of certain proteins crystallization was obtained in an opposite process. Vapor diffusion occurred from the reservoir towards the drop, leading to a more dilute salt solution in which the protein was less soluble [117]

2.2.2 Description of the diffraction by a crystal

To describe what happens when a plane wave is diffracted by a protein crystal, it is assumed that the crystal is small enough and that diffraction is weak enough so that the amplitude of the incident wave can be considered as constant in all parts of the crystal. This is called the kinematical approximation.
The amplitude of the scattered wave is then proportional to the Fourier transform of the electron density $\rho(\boldsymbol{r})$ in the crystal. As the ultimate goal is to obtain a model at atomic resolution, it is convenient to consider the electron density as an assembly of individual scatterers, i.e. the atoms in the structure. Therefore

$$\rho(\boldsymbol{r}) = \sum_i \boldsymbol{f}_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{2.1}$$

where $\boldsymbol{f}_i$ is the atomic form factor of the i-th atom and where $\boldsymbol{r}_i$ is its position vector. The structure factor $F(\boldsymbol{h})$ is the Fourier transform of $\rho(\boldsymbol{r})$. Additionally the movement of the atoms has to be taken into account. To do this the Debye-Waller factor or temperature factor W_i is introduced for every atom. With (2.1) $F(\boldsymbol{h})$ becomes:

$$F(\boldsymbol{h}) = \sum_i \boldsymbol{f}_i e^{(i2\pi \boldsymbol{r}_i \boldsymbol{h})} e^{-W_i} \tag{2.2}$$

By definition, a crystal is the periodic repetition of a unit cell on a lattice. Therefore the scattering factor can be written as the sum of the electron density $\rho_u(r)$ in the unit cell over all the nodes of the reciprocal lattice.

$$F(\boldsymbol{h}) = \sum_n \int_V \rho_u(\boldsymbol{r} + \boldsymbol{r}_n) e^{(i2\pi(\boldsymbol{r}+\boldsymbol{r}_n)\boldsymbol{h})} \mathrm{d}\boldsymbol{V} \tag{2.3}$$

Here r_n is a vector connecting two nodes of the reciprocal lattice. $F(\boldsymbol{h})$ takes only non-zero values, when $\boldsymbol{r}_n\boldsymbol{h}$ gives integer values.
If the amplitude and phase and therefore the structure factor of the scattered wave could be measured, then by a simple inverse Fourier transform, the electron density in the crystal could be calculated. However experimentally one can only measure intensities I, which are proportional to $|F(\boldsymbol{h})|^2$. This means that all information on the phase of the structure factor is lost. Yet the phase information is essential to reconstruct the electron density in the crystal. Therefore in order to solve a crystallographic structure the phase problem needs to be solved (see section 2.2.4).

2.2.3 Data acquisition and treatment

To collect a complete data set from a crystal, diffraction patterns are measured at different orientations of the crystal with respect to the beam. Usually the crystal is rotated by one degree between subsequent images. Also during the collection of one image the crystal is slowly rotated back and forth. This is to ensure that the nodes of the reciprocal lattice intersect with the Ewald sphere and thus that the diffracted waves interfere constructively.
Usually a few images are sufficient to determine the crystal lattice, unit cell parameters and the orientation of the crystal in the laboratory reference system from the pattern of the diffraction spots with the help of a computer program. This is always the first step in data treatment as it permits to define the most efficient collection strategy, i.e. how to get a complete data set while taking the fewest images possible. The next step in data processing is the integration of the images. This means that the peak profile for all observed spots is determined, which is not as straight forward as it may sound, because a single spot can be observed on several consecutive images. Once the peak profile is determined, its intensity may be estimated.

2.2.4 Solving the phase problem and refining the model

As mentioned in section 2.2.2 the amplitude but not the phase of the structure factor can be obtained from the diffracted intensities. Several methods have been developed to retrieve some information on the phase. The easiest one is molecular replacement. It requires the existence of an atomic model of a structure similar to the protein structure to be solved. To determine whether a structure could be sufficiently similar, the primary sequences are compared. When a suitable model is found, a rotation and a translation function are calculated that will place the model correctly in the new unit cell. Thereafter a Fourier transformation is applied to the model to obtain the

corresponding phases. Using the phases from the model and the measured amplitudes, an inverse Fourier transformation is carried out. As most of the structural information is contained in the phase, the electron density map calculated in this way has a strong bias towards the molecular replacement model. To eliminate this bias as much as possible and to compensate errors in the data, molecular replacement is followed by an iterative refinement process. Once a first electron density map has been calculated, the model that has been used for the phase information is placed in the electron density map. During the refinement process missing pieces of the amino acid chain are added, the amino acid sequence is corrected where the protein's sequence does not coincide with the model sequence and ligand and water molecules are added. In this step information that was not obtained in the crystallography experiment, but elsewhere, is used. This includes the primary sequence, information about bonding angles and preferred conformation of side chains as well as information about possible distances between neighboring atoms. Once the first model has been corrected, new phases are calculated from this model. These phases, together with the experimental amplitudes, serve to calculate a new electron density map. The model is then further refined with this new map. This process is repeated several times. Apart from visual inspection there are a number of calculation routines that can help to improve the model by minimizing the difference between the structure factors calculated from the model and those observed experimentally, while respecting stereochemical constraints. The problem with these methods is that there are a lot of local minima where the minimization can get stuck. Therefore it is very important to visually inspect the model between computational refinement cycles.
To estimate the quality of the model, the R-factor is calculated. It is defined as

$$R = \frac{\sum_n |F_{obs} - F_{calc}|}{\sum_n |F_{obs}|} \tag{2.4}$$

and the summation is carried out over all reflections. Similarly a second parameter R_{free} is defined. Here the summation is only carried out over a test set of reflections that have been designated from the beginning of refinement and that are not used in the minimization routines. As a rule of thumb at the end of the refinement R_{free} should be equal to the resolution in Ångström divided by 10.

2.3 Small angle scattering

Using small angle x-ray (SAXS) or neutron (SANS) scattering, low resolution structural information on biomolecules in solution can be obtained [118]. Whereas x-rays are sensitive to the electron density of the sample, neutrons interact with the nuclei. Small angle scattering (SAS) gives access to distances in particles that are large compared to interatomic distances. In a SAS experiment a solution of macromolecules in a buffer is placed into an x-ray or neutron beam and the scattering in the vicinity of the direct beam is recorded on a detector.
SAS experiments should be done on solutions that are ideal, i.e. sufficiently dilute to avoid interactions between the individual components, which can be individual molecules or complexes of several molecules. If this is the case and if the sample is monodisperse, i.e. made up of only one type of particles, structural information on the particles can be obtained. This includes the molecular weight M and the radius of gyration R_G as well as reconstructing the shape of the particle up to a resolution of about 10Å. In the case of ideal, polydisperse solutions average values for R_G and M can be calculated.

2.3.1 Theory

Assuming that the incident photons or neutrons do not exchange energy with the sample, only elastic scattering has to be considered. If $\mathbf{k}$ and $\mathbf{k_0}$ are the wave vectors of the scattered and of the incident beam, respectively, and the wave vector transfer is $\mathbf{q} = \mathbf{k} - \mathbf{k_0}$, then

$$q = |\mathbf{q}| = \frac{4\pi}{\lambda} sin\theta \tag{2.5}$$

where 2θ is the scattering angle and λ is the wavelength. Depending on whether neutrons or x-rays are used, the scattered intensity *in vacuo* corresponds to the Fourier Transform of the scattering length density of the nuclei ρ_n or of the electron density ρ_e. However, as the particles are in solution, ρ should be replaced by $\Delta\rho$, the contrast between the electron or scattering length density of the particle and that of the buffer solution, which is assumed to be homogeneous. In an ideal solution the particles are isotropically oriented and thus the scattered intensity is spherically symmetric around the beam. The total scattering intensity $I(q)$ can then be written as the sum of the scattering intensities of the individual components of the solution averaged over all possible orientations (denoted by $\langle ... \rangle$ in equation 2.6).

$$I(q) = \sum_n N_n \langle i_n(q) \rangle \tag{2.6}$$

N_n is herein the number of particles of type n and $i_n(q)$ is the intensity scattered by a particle of type n.

2.3.2 Data interpretation

2.3.2.1 Molecular weight

Jacrot and Zaccai [119] have shown that $I(0)$ is proportional to the concentration c and the molecular weight M of the different components in the solution provided the components have the same scattering density (i.e. the solution is monodisperse in composition)(see equation 2.7).

$$I(0) \propto \sum_n c_n M_n \tag{2.7}$$

As the intensity at $q = 0$ cannot be measured directly, because this region is masked by the beamstop, $I(0)$ has to be extrapolated from the data (see section 2.3.2.2). If the number of different particles is small and either their respective molecular weights M_n or their respective concentrations c_n are known, information about the unknown parameter M_n or c_n can be obtained from $I(0)$ using equation 2.7.

2.3.2.2 Radius of Gyration

The Guinier approximation [120] can be used to obtain further information on the particle in solution. Guinier has shown, that for very small angles, i.e. for $qR_G \leq 1$:

$$I(q) = I(0)exp\left(\frac{-q^2R_G^2}{3}\right) \tag{2.8}$$

where R_G is called the radius of gyration. R_G describes the distribution of scattering density around the center of scattering density of the particle. Spherical objects have smaller radii of gyration than elongated objects of the same mass. If $ln(I(q))$ is plotted against q^2, the data should yield a straight line with a slope of $R_G^2/3$ in the range of validity of the approximation. For ovoid shapes, such that of lysozyme, for example, the linear range is usually larger and can extend as far as $qR = 1.3$. By extrapolating the straight line to the ordinate $I(0)$ can be obtained.
If the solution is monodisperse ($n = 1$), the radius of gyration obtained from equation 2.8 can be associated with the shape and the size of the particles in solution. In the case of polydisperse systems the value for R_G will be influenced by all the different particles present yielding an average.

In the case of a monodisperse system it can be shown that:

$$I(q) = 4\pi \int_0^{D_{max}} P(r) \frac{sin(qr)}{qr} dr \tag{2.9}$$

The upper limit of the integral is D_{max}, the maximum distance between two points in the same particle. $P(r)$ is called the pair distribution function. Equation 2.10 shows how the radius of gyration is related to $P(r)$:

$$R_g^2 = \frac{\int r^2 P(r) dr}{2 \int P(r) dr} \tag{2.10}$$

This represents an alternative method to determine the radius of gyration.

2.3.2.3 Theoretical scattering curves

As the scattered intensity observed in SAS experiments is the rotationally averaged Fourier transform of the scattering length density ρ_n or of the electron density ρ_e of the particles in solution, theoretical scattering curves can be calculated if ρ_n or ρ_e are known. With the help of the programs CRYSOL (for x-rays) and CRYSON (for neutrons) that are part of the computer program package ATSAS [121], theoretical scattering curves can be calculated from pdb files. The obtained curves are very helpful for interpreting experimental scattering profiles or to check whether the structural model of the pdb file corresponds to the protein's solution structure.

2.4 Enzymatic activity

2.4.1 Enzymes

Enzymes are a class of proteins that catalyze chemical reactions in the organism. They speed up reaction rates by lowering the activation energy of the reaction. As all catalysts they are not altered themselves by the reaction, but they may undergo reversible changes. In some cases different protein molecules interact in order to form an enzymatically active protein complex. In order to fulfill their role in the organism enzymes, as all proteins, fold into a specific three-dimensional structure that is determined by their amino acid composition and is stabilized by weak interactions between different parts of the structure and between the structure and the solvent molecules. The region where the substrate is to bind to the enzyme is complementary to the substrate either in shape, charge or hydrophobicity, conferring great substrate specificity to the enzyme. During substrate binding, catalysis and

product release the enzyme may undergo important conformational changes such as the movement of different domains, subunits or elements of secondary structures. On a faster timescale sidechains or individual atoms will move. The dynamics of the enzyme are crucial for correct functioning. Some enzymes may also need to bind a small molecule or ion as a cofactor to work optimally (or to work at all).

2.4.2 Michaelis-Menten kinetics

To investigate how enzymes work, activity tests are carried out by measuring the product release of the enzyme as a function of time. The easiest model to describe enzyme kinetics was proposed by Michaelis and Menten in 1913 [122]. In their model the enzyme E binds reversibly to the substrate S. An intermediate complex ES is formed which can then react irreversibly to generate the product P and free enzyme.

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P$$

k_{off}, k_{on} and k_{cat} are the rate constants of the corresponding reaction. An important assumption when deriving the Michaelis-Menten equation is that after a very short initial phase the concentration of ES is constant, i.e.

$$\frac{\mathrm{d}[ES]}{\mathrm{d}t} = 0$$

The maximum reaction velocity v_{max} is reached when there is an excess of substrate. In this case the entire free enzyme is immediately bound to form the complex ES. Therefore

$$v_{max} = k_{cat}[E_0]$$

when $[E_0]$ is the total enzyme concentration.

The Michaelis-Menten constant is defined as:

$$K_m = \frac{k_{off} + k_{cat}}{k_{on}}$$

If k_{cat} is much smaller than k_{off},

$$K_m \approx \frac{k_{off}}{k_{on}} = K_s$$

K_s is the dissociation constant of the Michaelis-Menten complex $[ES]$. In this case K_m is a measure of the affinity of the enzyme for its substrate. The

Michaelis-Menten model can also be applied to enzymatic reactions that involve several steps. In the case when the Michaelis-Menten complex undergoes a two-step reaction, like it is observed for some proteases (e.g. TET) the reaction scheme can be written as follows:

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ES \xrightarrow[-P_1]{k_2} EA \xrightarrow[+H_2O]{k_3} E + P_2$$

In this case the Michaelis-Menten constant becomes

$$K_m = K_S \frac{k_3}{k_2 + k_3}$$

and k_{cat} depends on the two rate constants k_2 and k_3:

$$k_{cat} = \frac{k_2 k_3}{(k_2 + k_3)}$$

If k_2 and k_3 are very different in magnitude, k_{cat} will exclusively depend on the smaller rate constant. The reaction step associated with the smaller rate constant will be called the rate-limiting step of the reaction.
It can be shown that the velocity of product formation is given by:

$$v = \frac{\mathrm{d}[P]}{\mathrm{d}t} = v_{max} \frac{[S]}{K_m + [S]} \tag{2.11}$$

This equation is only valid if substrate and enzyme can diffuse freely. Inside cells the high macromolecular concentration can limit free diffusion. Also in the case of membrane proteins free diffusion is limited. In these cases other descriptions have to be used.
When plotting $\frac{\mathrm{d}[P]}{\mathrm{d}t}$ as a function of $[S]$ the graph shown in figure 2.1 is obtained. As the substrate concentration increases, the initial velocity of the reaction increases as it approaches v_{max} asymptotically. When $[S] = K_m$, $v = \frac{1}{2} v_{max}$. Therefore by experimentally determining the graph $v([S])$ the two parameters describing the reaction kinetics, v_{max} and K_m can be obtained. This is done either by fitting the data by non-linear regression or by representing the data in a way to obtain a straight line, e.g. in a Hanes-Woolf plot, which allows the use of linear regression methods.

In order to measure $v([S])$ a series of experiments with constant enzyme concentration and different substrate concentrations is carried out. For each substrate concentration the amount of product in the solution as a function of time is measured. Initially the product concentration increases linearly until the substrate concentration starts to decline. At this point the speed of the reaction decreases and approaches a plateau when all substrate is used.

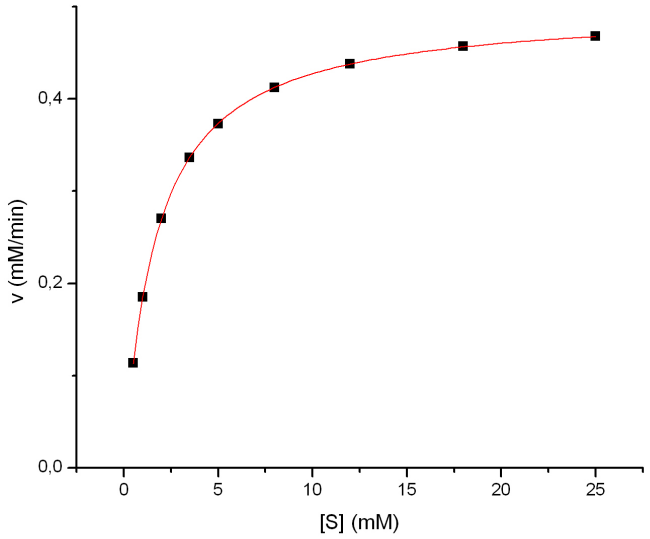


Figure 2.1: Reaction velocity v as a function of substrate concentration [S] for an enzyme following Michaelis-Menten kinetics.

To determine v at a given substrate concentration, a straight line is fitted to the first (linear) part of the curve.
Determining the parameters v_{max} and K_m under different conditions can give insight into how the catalytic process is influenced by these conditions. v_{max} is the maximum catalysis rate the enzyme can achieve, when all binding sites are continuously saturated with substrate. As K_m is the concentration of substrate at which the reaction rate v is half the maximum rate, K_m allows to estimate at which substrate concentration the reaction rate will be very close to v_{max}.

2.5 Analytical ultracentrifugation

In an analytical ultracentrifugation (AUC) experiment the sedimentation of the particles in the centrifugal field is analyzed. There are two types of experiments in AUC, sedimentation velocity and sedimentation equilibrium. Here only sedimentation velocity experiments are described.

2.5.1 Experimental setup

To monitor the sedimentation of biological macromolecules in a centrifugal field, the macromolecular concentration in the sample is monitored as a function of the radial position r. There are two different techniques to measure the concentration in the centrifugation cell. One is to follow the absorption of the sample at a chosen wavelength λ, typically 280 nm for proteins. The other is to follow the interference of the two parts of a split coherent beam. One part of the beam passes through a compartment of the centrifugation cell that contains the sample in its buffer, whereas the other part passes through a compartment that contains only the buffer. The difference of the refractive index between the two compartments is directly linked to a difference in concentration and can be seen as a shift of the pattern of the interference fringes.

2.5.2 Describing sedimentation

The sedimentation of a particle that is subjected to the centrifugal acceleration $\omega^2 r$ and sediments with a radial velocity u can be described by its sedimentation coefficient s, which is given by the Svedberg equation [123]:

$$s = \frac{u}{\omega^2 r} = \frac{M(1 - \overline{v}\rho)}{N_A f} \tag{2.12}$$

where M is the molar mass of the sedimenting particle, $\overline{v}$ is its partial specific volume, ρ is the density of the solvent, f is the frictional coefficient and N_A is Avogadro's number. From Stoke's law the frictional coefficient f_0 for a sphere can be derived:

$$f_0 = 6\pi\eta R_0$$

In this equation η is the viscosity of the solution and R_0 is the radius of the sphere. The frictional coefficient f for particles of other shapes is related to the hydrodynamic radius R_H:

$$f = 6\pi\eta R_H$$

The frictional ratio f/f_0 of a particle characterizes the deviation of its shape from a sphere. By inserting the expression for f_0 in Svedberg's equation (2.12), the maximum sedimentation coefficient s_{max} for a particle of given molar mass can be calculated. By comparing the experimental s to s_{max}, the frictional ratio can be determined and some information about the molecular shape can be deduced.

2.5.3 Data analysis

In an AUC experiment concentration profiles in a radial direction are measured as a function of time. To analyze this data, Lamm's equation [124] is used, which describes the concentration distribution χ

$$\frac{\partial \chi(r,t)}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}[rD\frac{\partial \chi(r,t)}{\partial r} - s\omega^2 r^2 \chi(r,t)]$$

where D is the diffusion coefficient

$$D = \frac{RT}{N_A f}$$

and R is the gas constant. Computer programs are used to fit the measured data with Lamm's equation. In a usual AUC experiment there is more than one species of molecules sedimenting. Therefore a distribution of sedimentation coefficients is assumed. If $a(r,t)$ are the experimental data and $c(s)$ is the concentration of the species with sedimentation coefficient between s and $s+ds$ then

$$a(r,t) = \int c(s)\chi(s, D(s), r, t)ds + \epsilon$$

When fitting this equation to the data, an average frictional ratio is used for all sedimenting species to establish $D(s)$ [125]. The experimental error is accounted for by the introduction of the parameter ϵ.
Once the experimental sedimentation coefficients have been determined, they can be compared to the theoretical values calculated with Svedberg's equation (2.12). One possible application of this is to determine the oligomeric state of the protein as different oligomers will have different molar masses. In order to compare sedimentation coefficients obtained from experiments in different solvents, a normalized sedimentation coefficient $s_{20,w}$ is calculated. This corrected sedimentation coefficient reflects how the particle would sediment in water at 20 °C. The relation between the experimental sedimentation coefficient s and $s_{20,w}$ is shown in equation 2.13.

$$s_{20,w} = s\frac{(1-\rho_{20,w}\overline{v})}{(1-\rho\overline{v})}\frac{\eta}{\eta_{20,w}} \tag{2.13}$$

In this equation ρ and η are the density and the viscosity of the solvent used in the experiment and $\rho_{20,w}$ and $\eta_{20,w}$ are those of water at 20 °C.

Chapter 3

Characterization of PhTET3

3.1 Introduction

In the genome of *Pyrococcus horikoshii* three genes encoding TET proteases have been identified. This system of homologous enzymes from the same organism represents an ideal model to study the structural basis for enzyme specialization. The structures of PhTET1 and PhTET2 have been solved [107, 108, 103]. In this publication we present the x-ray crystallographic structure, the biochemical characterization and some biophysical data on PhTET3. By combining these results with information that had been published before, we were able to specify the mode of action of the TET proteases and to reveal that the three TET proteins are not functionally redundant in *P. horikoshii*.

In our work we have used biochemical and biophysical techniques and the combination of the data has allowed to obtain complementary information. Therefore we decided that it would be appropriate to have two first authors for this work. Asunción Durá has done the biochemical work and has taught me how to purify the protein. My contribution was the determination of the protein's structure and the SANS studies. We analyzed the structure together, which turned out to be very profitable as Asun contributed from the biochemical perspective and I amended the biophysical point of view. Hence this publication is a genuinely interdisciplinary work.

mmi_6600

Molecular Microbiology (2009) ■ doi:10.1111/j.1365-2958.2009.06600.x

The structural and biochemical characterizations of a novel TET peptidase complex from *Pyrococcus horikoshii* reveal an integrated peptide degradation system in hyperthermophilic Archaea

M. Asunción Durá,[§] Eva Rosenbaum,[§] Amédé Larabi,[†] Frank Gabel, Frédéric M. D. Vellieux and Bruno Franzetti*

Institut de Biologie Structurale J.-P. Ebel, UMR 5075 CNRS-CEA-UJF, 41 rue Jules Horowitz, 38027 Grenoble, France.

Summary

The structure of a 468 kDa peptidase complex from the hyperthermophile *Pyrococcus horikoshii* has been solved at 1.9 Å resolution. The monomer contains the M42 peptidase typical catalytic domain, and a dimerization domain that allows the formation of dimers that assemble as a 12-subunit self-compartmentalized tetrahedron, similar to those described for the TET peptidases. The biochemical analysis shows that the enzyme is cobalt-activated and cleaves peptides by a non-processive mechanism. Consequently, this protein represents the third TET peptidase complex described in *P. horikoshii*, thereby called PhTET3. It is a lysyl aminopeptidase with a strong preference for basic residues, which are poorly cleaved by PhTET1 and PhTET2. The structural analysis of PhTET3 and its comparison with PhTET1 and PhTET2 unravels common features explaining the general mode of action of the TET molecular machines as well as differences that can be associated with strong substrate discriminations. The question of the stability of the TET assemblies under extreme temperatures has been addressed. PhTET3 displays its maximal activity at 95°C and small-angle neutron scattering experiments at 90°C demonstrate the absence of quaternary structure alterations after extensive incubation times. In conclusion, PhTETs are complementary peptide destruction machines that may play an important role in the metabolism of *P. horikoshii*.

Accepted 31 December, 2008. *For correspondence. E-mail franzetti@ibs.fr; Tel. (+33) 438 78 95 69; Fax (+33) 438 78 54 94. †Present address: EMBL Grenoble, BP 181, 6 rue Jules Horowitz, 38042 Grenoble Cedex 9, France. §These authors have contributed equally to this work.

Introduction

In all cell types, cytosolic proteases recognize and break down misfolded and malfunctioning proteins that are produced as a result of environmental stress, mutations or errors in biosynthetic processes. Misfolded polypeptides are prone to aggregation, and consequently, must be scavenged and degraded (Baumeister *et al.*, 1998; Ward *et al.*, 2002). Moreover, basal levels of regulatory proteins must be kept low by rapidly removing these proteins when they are no longer needed (Gottesman, 2003). Therefore, intracellular protein degradation plays a central role, not only in maintaining the integrity of the proteome under harsh environmental conditions, but also for the regulation of crucial events such as the cell cycle, metabolic switches and signal transduction (Kirschner, 1999). Hyperthermophilic microorganisms that thrive close to the deep-sea hydrothermal vents are exposed to sharp gradients of different physicochemical factors such as temperature, pressure and nutrient availability. This may greatly influence their metabolisms and the folded state of their proteins. In this context, it is generally assumed that proteolysis represents an important aspect of microbial adaptation to extreme environments (Ward *et al.*, 2002). Moreover, many hyperthermophilic Archaea derive energy primarily from the degradation of external complex proteinaceous substrates and it has been hypothesized that a high level of proteinase activity is necessary for their growth (Snowden *et al.*, 1992; Godfroy *et al.*, 2000).

Intracellular proteolysis requires very intricate biochemical machinery (Kirschner, 1999) and is subject to spatial and temporal control in order to prevent the damage of proteins not destined for destruction (Baumeister *et al.*, 1998). Proteolytic degradation is a multistep process involving initial digestion of the target protein to peptide fragments that are further broken down to smaller peptides and amino acids (Tomkinson, 1999). Different proteases are implicated in the process. These

enzymes range from simple monomeric hydrolases to complex multisubunit structures with molecular masses in the order of 1 MDa (Ward *et al.*, 2002). Some of these high-molecular-weight proteases are self-compartmentalized; that is, they form cellular sub-compartments through self-association and enclose inner cavities that harbour the proteolytic active sites. These are only accessible to unfolded polypeptides and, in this way, cytoplasmic proteins are protected from unwanted cleavage (Tamura *et al.*, 1996; Lupas *et al.*, 1997). Although few in number, self-compartmentalized proteases represent the main agents of intracellular protein breakdown (Baumeister *et al.*, 1998).

In Eukarya and Archaea, the core of the proteolytic system is the 20S proteasome, a hollow barrel-shaped multimer that carries out the first degradation step (Baumeister *et al.*, 1998). The 20S proteasome is a self-compartmentalized non-specific endoprotease that processively cleaves cell proteins to peptides 2–30 residue long (Kisselev *et al.*, 1998; 1999; Saric *et al.*, 2004). The proteins to be degraded are previously unfolded and translocated into the proteolytic complex by ATP-dependent regulatory particles (Sauer *et al.*, 2004). The further hydrolysis of the peptides generated by the proteasome requires energy-independent metallopeptidases (Saric *et al.*, 2004). Degradation of most proteasome products is initiated by endoproteolytic cleavages performed by oligopeptidases, and the resulting six to nine residue fragments are further digested by tripeptidyl- and dipeptidyl-peptidases, and finally by aminopeptidases, carboxypeptidases and di- and tripeptidases to yield free amino acids (Gonzales and Robert-Baudouy, 1996; Tamura *et al.*, 1998; Tomkinson, 1999; Saric *et al.*, 2004).

In the archaeon *Thermoplasma acidophilum*, the tricorn endopeptidase (TRI) has been proposed to hydrolyse the proteasome products to smaller fragments (Tamura *et al.*, 1996). TRI forms a hexameric ring assembly and when it interacts with several smaller aminopeptidases, the resulting complex degrades oligopeptides in a sequential manner, yielding free amino acids (Tamura *et al.*, 1996; 1998). Most archaea lack TRI homologues. Instead, they contain homologues of TET, a dodecameric aminopeptidase complex with a tetrahedral shape (Franzetti *et al.*, 2002). TET belongs to the metallopeptidase family M42 according to the MEROPS peptidase classification (Rawlings *et al.*, 2008) and can process peptides that are more than 25 amino acids in length (Franzetti *et al.*, 2002; Durá *et al.*, 2005). The TET interior is accessible through eight openings situated in the facets and apices of the assembly and its architecture is different from all the proteolytic complexes described to date, most of which are made up by rings or barrels with a single central channel and only two openings (Franzetti *et al.*, 2002; Russo and Baumann, 2004; Borissenko and Groll, 2005; Schoehn *et al.*, 2006).

Pyrococcus horikoshii is an anaerobic archaeon, which grows optimally at 98°C. It was isolated from hydrothermal fluid samples obtained at a depth of 1395 m in the Pacific Ocean (González *et al.*, 1998). While the genomes of halophilic and methanogenic archaeons contain only one type of TET protein, three open reading frames (ORFs) encoding for TET-homologous proteins were detected in the genome of *P. horikoshii*: PH0519, PH1527 and PH1821. The proteins encoded by these ORFs share more than 64% of similarity in their amino acid sequence. The PH0519 ORF encodes a thermostable and cobalt-activated aminopeptidase described first by Ando *et al.* (1999). The 3D structures of PH0519 show that it can assemble alternatively as a tetrahedral dodecameric particle (therefore called PhTET1-12s) or as an octahedral tetracosameric edifice (PhTET1-24s) (Schoehn *et al.*, 2006). The PH1527 product, named PhTET2, was found to be a 12-subunit tetrahedral aminopeptidase that displays a non-processive catalytic activity against N-terminal aliphatic and neutral amino acids (Durá *et al.*, 2005). As for PhTET1, the X-ray structure of the PhTET2 edifice revealed the self-compartmentalizing nature of the complex (Russo and Baumann, 2004; Borissenko and Groll, 2005).

In this communication we present the X-ray structure of the 468 kDa dodecameric assembly formed by the PH1821 protein together with its enzymatic characterization and the study of its thermal stability by small-angle neutron scattering (SANS). The complex has a tetrahedral shape and therefore, it has been named PhTET3. The structural and biochemical properties of PhTET3 are discussed and compared with those of PhTET1 and PhTET2. The results of this analysis allow specifying the mode of action of the TET proteases and reveal that the three TET proteins are not functionally redundant in *P. horikoshii* and probably in the other hyperthermophilic Archaea.

Results

PhTET3 purification

PhTET3 was purified as described in the *Experimental procedures* section. The absorption at 280 nm on the chromatogram of the gel filtration step displayed a symmetric peak. According to a prior calibration of the column, this peak corresponds to a molecular mass of *c.* 500 kDa. In addition, the lysyl aminopeptidase activity measured in the fractions obtained from the peak followed its shape. A sample from the pooled fractions was run on a SDS-PAGE gel and gave a single band of a molecular weight of approximately 40 kDa. Together with the results from the gel filtration, this showed that the protein was pure and forming a dodecameric complex. This was further con-

firmed in electron micrographs, where the quaternary structure of the purified protein appeared in negative stain as a homogeneous population of tetrahedral particles (not shown).

Topology of the PhTET3 subunit

The 12-subunit PhTET3 complex was crystallized and its 3D structure was solved at a resolution of 1.9 Å by molecular replacement using the 1.6 Å resolution structure of the monomer of PhTET2 as the search probe (Protein Data Bank code 1Y0Y). The asymmetric unit of the crystal contains one monomer of ~39 kDa. No electron density could be observed for the first eight N-terminal residues as well as for residues 128–136. The monomer is wedge-shaped with dimensions ~72 × 48 Å. It contains two domains, a larger catalytic domain (residues 9–75 and 169–354) and a smaller dimerization domain (residues 76–168).

The catalytic domain is a flattened globule made up of a central mixed eight-stranded β-sheet surrounded by seven α-helices (Fig. 1A, orange). It comprises an additional motif of a mixed four-stranded β-sheet near the dimerization domain (Fig. 1A, magenta). Two regions of large positive electron density in the active site were modelled as two divalent zinc ions (Fig. 1A, green) in agreement with an X-ray fluorescence scan performed on the crystal that showed a zinc absorption line. The metal ligands (Fig. S1) are identical to those present in PhTET1 (Schoehn *et al*., 2006), PhTET2 (Russo and Baumann, 2004; Borissenko and Groll, 2005) and the *Aeromonas* aminopeptidase (Chevrier *et al*., 1994), the type example of clan MH, in which family M42 is included (Rawlings *et al*., 2008). Moreover, the type of co-ordination is identical in all four proteins. Therefore, the catalytic mechanism of hydrolysis by *Aeromonas* aminopeptidase and PhTETs should be extremely similar, if not identical (Stamper *et al*., 2001; Russo and Baumann, 2004; Borissenko and Groll, 2005; Schoehn *et al*., 2006).

The dimerization domain has β-barrel topology and a cylindrical shape. It mainly consists of a mixed six-stranded β-sheet surrounded by three α-helices at the small ends (Fig. 1A, blue). It also contains a flexible loop sticking out at one end and made up of residues 121–138. No electron density could be observed for residues 128–136.

PhTET3 assembles as a tetrahedral edifice

By applying the crystallographic symmetries of the I23 space group to the PhTET3 monomer, we solved the structure of a dodecameric complex (Fig. 1C), whose dimensions are in agreement with the SANS data obtained for the protein in solution (see below). Therefore, we conclude that the quaternary structure of PhTET3 is similar to that reported for PhTET1 and PhTET2 (Russo and Baumann, 2004; Borissenko and Groll, 2005; Schoehn *et al*., 2006). All three proteins form tetrahedral particles composed of 12 monomers. The building blocks of the tetrahedra are most likely dimers that form the six edges of the particle (Schoehn *et al*., 2006) (Fig. 1C, left). The analysis of the dimer–dimer interface shows that the interactions between dimers are located in three distinct regions of the interface (Fig. S2). On one edge of the interface (Fig. S2, left), a salt bridge between K51 and E232 is observed. On the other end (Fig. S2, right), a cluster of salt bridges and hydrogen bonds involving residues N80, D140, K248 and K250 is found. Residues Q214, R220, P265, N300 and Q303 form a second cluster of hydrogen bonds in the centre of the interface. No hydrophobic interactions are detected.

On the vertices of the PhTET dodecamers, three monomers are arranged in a triangular fashion, forming a pore in the middle of the triangle (Russo and Baumann, 2004; Borissenko and Groll, 2005; Schoehn *et al*., 2006) (Fig. 1C, middle). In PhTET3, these pores are blocked by the Y227 residues from each monomer, with a distance between successive hydroxyl groups of only 3.5 Å. Furthermore, residues K306 and R223 stick out into the pore. The faces of the PhTET tetrahedra are formed by three dimers that are also arranged in a triangle, forming a larger opening in the middle of it (Russo and Baumann, 2004; Borissenko and Groll, 2005; Schoehn *et al*., 2006) (Fig. 1C, right). This opening is partially blocked by three K84 residues in the PhTET3 edifice, as the distance between successive ε-amino groups is 8.4 Å. However, if an alternate conformer of the side-chains of these three lysyl residues is taken (without provoking steric clashes with the remainder of the polypeptide), that distance is increased to 15.7 Å.

The active sites of the PhTET complexes are located inside the particles, in the catalytic chambers that are enclosed by the tetrahedron apices (Fig. 2). The walls of these chambers are positively charged (Fig. 2B), and each one contains three active sites arranged in a circular way. Adjacent to the active sites, the specificity pockets appear as small cavities (Fig. 2A, detail).

PhTETs structural comparison

The secondary and tertiary structures of the PhTET3 subunit are similar but not identical to those reported for PhTET1 (Schoehn *et al*., 2006) and PhTET2 (Russo and Baumann, 2004; Borissenko and Groll, 2005). Even though the PhTET1 monomer contains 22 residues less than the PhTET3 monomer (Fig. 1B), the two proteins contain roughly the same secondary structure elements. In PhTET3, several of these elements are longer than

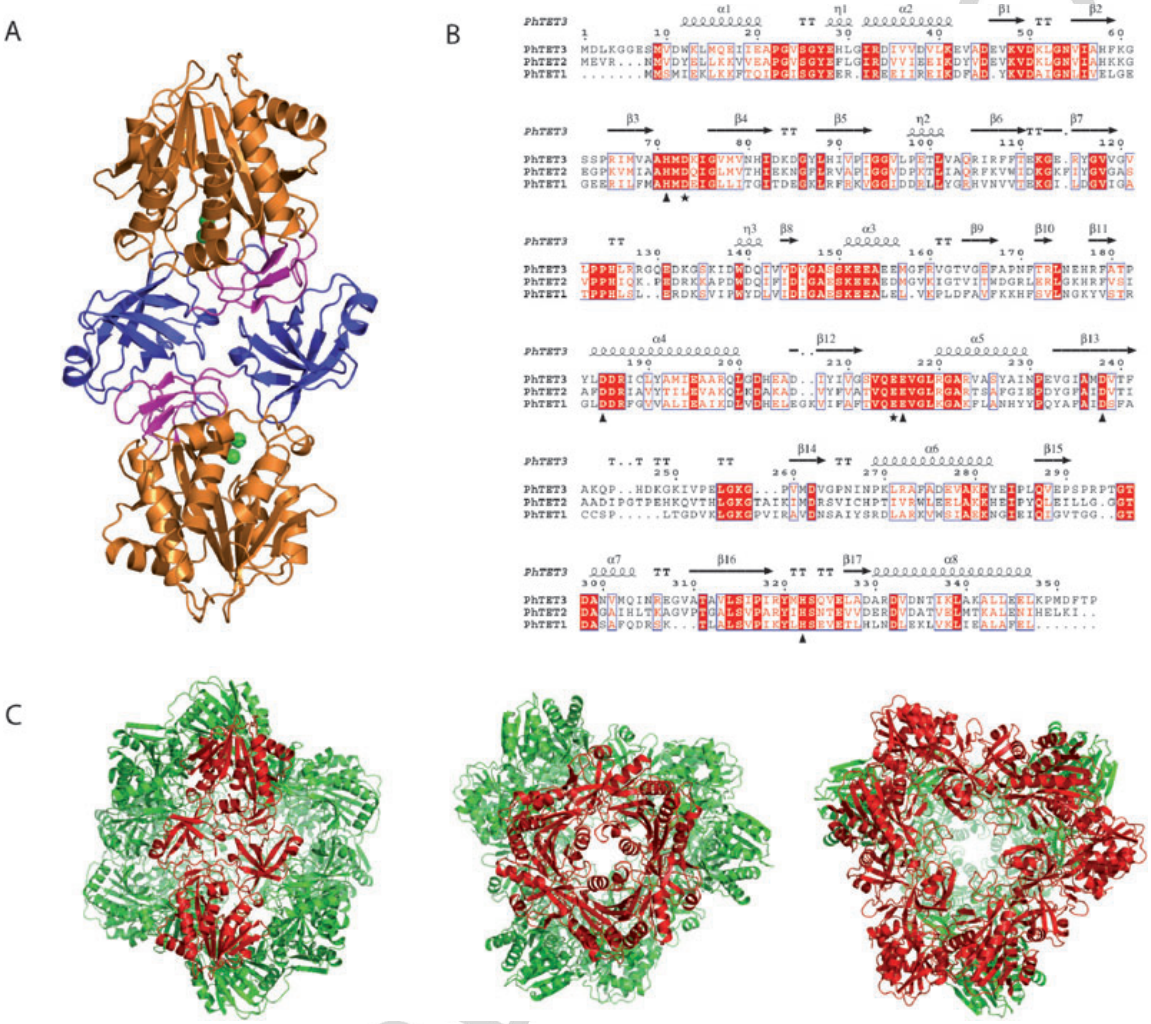


Fig. 1. A. View of the PhTET3 dimer along the twofold symmetry axis. The catalytic domain of each monomer is coloured in orange with the additional four-stranded β-sheet in magenta. The dimerization domain is represented in blue and the catalytic zinc ions are depicted as green spheres. The dimerization domain of each monomer interacts with the peripheral β-sheet of the catalytic domain of the other monomer to form the dimer.
B. Multiple sequence alignment of the three TET proteins from *P. horikoshii* with the secondary structure elements of PhTET3 as determined by DSSP (Kabsch and Sander, 1983). Residues involved in the binding of the catalytic metal ions are marked with a triangle; catalytic residues are marked with a star. These residues are conserved between the three proteins.
C. Organization of the PhTET3 complex. On the left, the complex is viewed down a twofold symmetry axis. One of the edges of the tetrahedron, formed by a dimer, is highlighted in red. In the middle, a view down a threefold axis with an apex of the tetrahedron pointing upwards is depicted. The three monomers forming the apex are coloured red. On the right, the image shows a view down a threefold axis but in the opposite direction so that one face of the tetrahedron, formed by three dimers (red), is visible.

their counterparts in PhTET1, especially the C-terminal α-helix, which is formed by 17 residues in PhTET3, but only by eight residues in PhTET1. The PhTET2 monomer has one residue less than PhTET3, but shows several structural elements that are absent in PhTET3. Three 3_{10} helices (η2, η5 and η6, following Russo and Baumann (2004) assignment) from PhTET2 are not present in PhTET3 as is one β-stand, β16. At its C-terminus, PhTET2 has an α-helix (α8) followed by a 3_{10} helix (η6). Only the last three C-terminal residues of PhTET2 are without secondary structure, whereas in PhTET3 the α-helix (α8) is followed by eight residues that do not form any secondary structure (Fig. 1B).

Regarding the quaternary structure, in the PhTET1 complex the openings on the faces of the tetrahedron are larger than those in PhTET2 and PhTET3, where the

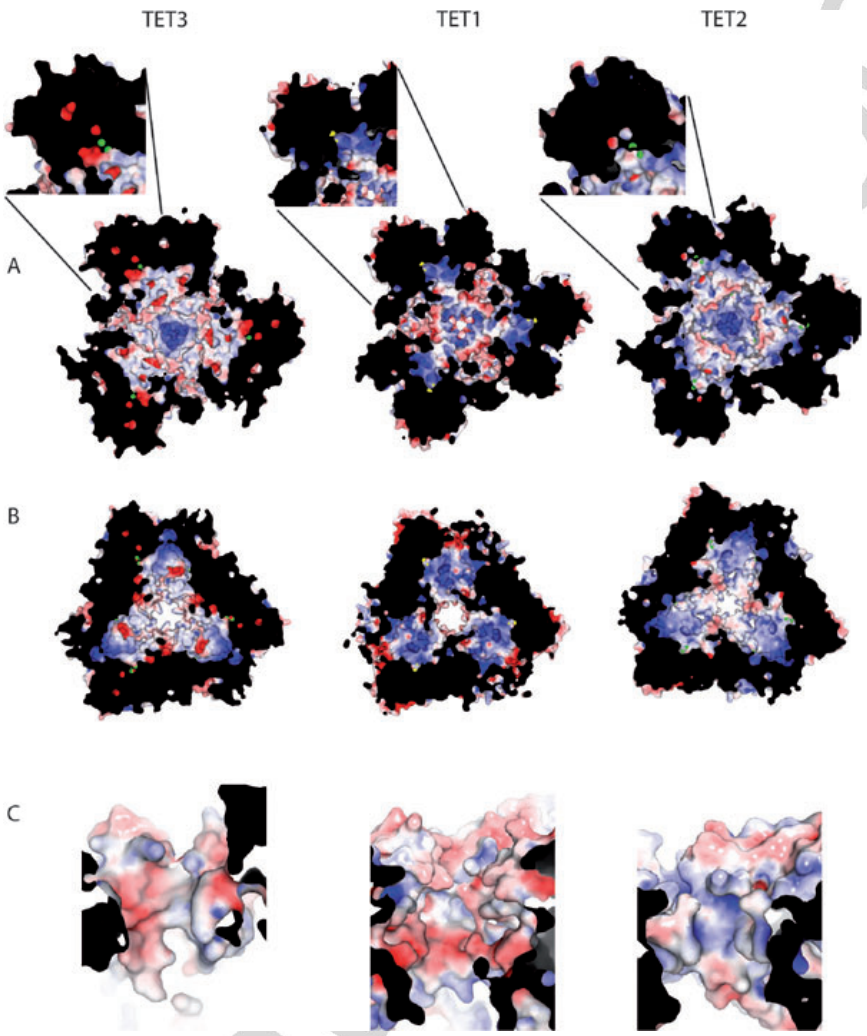


Fig. 2. A. Cut-open surface representation of the *P. horikoshii* TET complexes viewed down one of the large pores located in the faces of the tetrahedra. The catalytic sites and their surroundings are shown in detail.
B. View of the same complexes in the opposite orientation.
C. Detailed depiction of the channel walls next to the face pores of the tetrahedra. The lower part of the channel opens to the internal cavities of the complexes. The surfaces are coloured according to their solvent accessible electrostatic surface from –8 kT e^{-1} (red) to +8 kT e^{-1} (blue). Calculations were performed at 363 K, pH 7 with 150 mM ionic strength, a solute dielectric of 2 and a solvent dielectric of 80. The catalytic ions are coloured yellow (Co^{2+}) and green (Zn^{2+}).

entrance is partially blocked by K81 and K84 respectively (Fig. 2B). As the pores are between three subunits, all residues are repeated three times in a circular fashion around the holes' walls. Furthermore, in PhTET1 a small opening can be seen in the apex, which has been suggested to serve for the expulsion of the free amino acids produced after catalysis (Schoehn *et al.*, 2006). However, the apex of PhTET2 and PhTET3 is blocked by F224 and Y227 respectively. The presence of K307 and R220 in PhTET2, and K306 and R223 in PhTET3 gives the whole apex region a very positive charge. In PhTET1 the side-chains of three D291 residues form a layer of negative charges (Fig. 2B).

The active site residues are conserved in all three particles (Fig. 1B). However, the electrostatic charge distribution in the specificity pockets is very different. Whereas in PhTET1 the presence of R120, H115, N252 and K212 gives a very positive charge, in PhTET2 the pocket adjacent to the active site is essentially uncharged due to the presence of the side-chains of L216, V236, I238, L293

and I322. Finally, in PhTET3, the specificity pocket is formed by the side-chains of D262, T240, T295 and T297, which leads to a negative charge (Fig. 2A, detail). In addition, the geometry of these pockets is different in each PhTET complex, as well as their location in the particles relative to the active sites.

Figure 2C shows the electrostatic charge distribution in the channel that leads to the interior of the PhTETs from the large pores. Note that in PhTET1 a well-defined entrance channel can be seen, whereas in PhTET3 and PhTET2 the channel is shorter, and can be considered as a pore that opens to the big central cavity. In the three PhTET particles, consecutive layers of different electrostatic charges can be distinguished. For PhTET3, a negatively charged layer is found next to the pore (Fig. 2C, left). This layer is formed by D83, D138 and D140 and is followed by a layer of positive charge due to the presence of K84, K151, K205, K243, H246, K248 and R319. The most internal negatively charged layer is formed by D85, E152 and E153. In PhTET2, the external layer is negatively charged as well and formed by the residues D135 and D137. The positively charged middle layer is made up of H78, K81, K148, H248 and H253, and it is followed by a third layer formed by E146, E149, E150, E152 and D241. The electrostatic charge distribution in the channels of PhTET1 has been described by Schoehn *et al.* (2006) and consists also of three layers of opposing charges, with the exception of that located in the entrance that is mixed, as it comprises the side-chains of D74, E75 and K240. Interestingly, only a small number of interactions between residues from different layers are detected, even though these residues have different charges.

The dodecameric PhTET3 complex is a lysyl aminopeptidase

The cleavage specificity of PhTET3 was studied first by using different chromogenic [4-nitroaniline (pNA) conjugated] and fluorogenic [7-amino-4-methylcoumarin (AMC) conjugated] aminoacyl compounds. As shown in Table 1, the enzyme acted only on a limited number of these aminopeptidase substrates. Optimal amidolytic activity was observed against Lys-pNA. Other substrates also hydrolysed but at lower rates were Arg-, Glu- and Leu-pNA. Poor activity was detected against ornithinyl- and Gln-AMC, and Asp-, Met- and Ala-pNA, while the enzyme showed very poor or no activity against the other aminoacyl-pNA or aminoacyl-AMC assayed.

PhTET3 is a strict aminopeptidase devoid of N-terminal deblocking activity

Exopeptidase activity and specificity depend on the sequence composition and length of the target peptide

Table 1. Activity of PhTET3 on different synthetic chromogenic and fluorogenic compounds.

Substrate	Concentration (mM)	Relative activity[a] (%)
Aminoacyl-pNAs		
Lys-pNA	5	100
Arg-pNA	5	66.6
Glu-pNA	5	20.1
Leu-pNA	5	11.4
Asp-pNA	5	3.0
Met-pNA	5	2.6
Ala-pNA	5	2.0
Gly-pNA	5	0.6
Phe-pNA	2	0.3
Pro-pNA	5	0
His-pNA	5	0
Ile-pNA	5	0
Acetyl-Leu-pNA	2	0
Pyroglutamyl-pNA	5	0
Aminoacyl-AMCs		
Ornithinyl-AMC	0.5	6.5
Gln-AMC	0.5	5.5
Ser-AMC	0.5	0.5
Asn-AMC	0.5	0.4
Tyr-AMC	0.5	0.3
Trp-AMC	0.5	0.0
Val-AMC	0.5	0.0
Thr-AMC	0.5	0
Peptides		
Ala-Ala-pNA	5	0.9[b]
Ala-Ala-Ala-pNA	5	0.3[b]
Gly-Pro-pNA	5	0.2
Ala-Pro-pNA	5	0.2
Benzyloxycarbonyl-Gly-Gly-Leu-pNA	1	0.1
Succinyl-Ala-Ala-Phe-pNA	5	0.0
Succinyl-Leu-Tyr-AMC	0.5	0.0
Succinyl-Ala-Ala-pNA	5	0
Ala-Ala-Phe-AMC	0.5	0
Benzyloxycarbonyl-Gly-Gly-Arg-AMC	0.5	0
Benzyloxycarbonyl-Leu-Leu-Glu-β-naphthylamine	0.5	0
Succinyl-Leu-Leu-Val-Tyr-AMC	0.5	0

a. Expressed as a percentage of the activity against Lys-pNA, which was given a value of 100%.
b. Apparent relative activity (see text).

and some aminopeptidases have been shown to exhibit also endopeptidase or di-/tripeptidyl peptidase activity that can only be detected by studying the peptide degradation process (Singh and Kalnitsky, 1980; Taylor, 1993; Koldamova *et al.*, 1998; Geier *et al.*, 1999; Turk *et al.*, 2001; Chandu *et al.*, 2003). Consequently, PhTET3 activity was assayed by using various chromogenic and fluorogenic peptides. No activity could be measured against any of the N-terminal blocked derivatives (Table 1). However, we observed that PhTET3 generated pNA from Ala-Ala-pNA and Ala-Ala-Ala-pNA, although very slowly (Table 1). In order to determine if the detected activity was the consequence of a di-/tripeptidyl action or was the result of sequential aminopeptidase cuts, we investigated

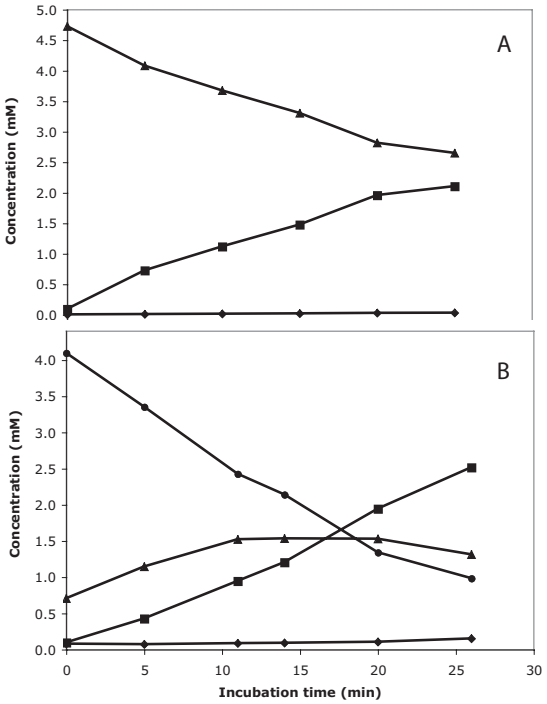


Fig. 3. A. Ala-Ala-pNA degradation by PhTET3.
B. Ala-Ala-Ala-pNA degradation by PhTET3. Concentration evolution of the substrates and hydrolysis products in the reaction mixture (●, Ala-Ala-Ala-pNA; ▲, Ala-Ala-pNA; ■, Ala-pNA; ◆, pNA).

the peptidyl- and aminoacyl-pNA composition of the reaction mixtures at different time points by reversed-phase HPLC. When Ala-Ala-pNA was used as a substrate, Ala-pNA accumulated very fast in the reaction mixture, and the Ala-pNA/pNA ratio diminished from 123 at 5 min to 69 at 25 min (Fig. 3A), therefore indicating that the enzyme was mostly cleaving the substrate to Ala-pNA, and then also this last product to give pNA. In the same way, we studied the products of the PhTET3-driven hydrolysis of Ala-Ala-Ala-pNA. In this case, Ala-Ala-pNA was produced, but also consumed by the enzyme (Fig. 3B) while pNA was generated after a lag phase. At 26 min the ratios Ala-Ala-pNA/pNA and Ala-pNA/pNA were 8 and 33 respectively. These results agree with strict aminopeptidase behaviour of PhTET3.

PhTET3 is a cobalt-activated enzyme

The effect of several metal ions on PhTET3 enzyme activity is shown in Table 2. Co^{2+} ions had a stimulatory effect on the amidolytic activity while the rest of the cations assayed caused a clear inhibition at 1 mM, especially Cu^{2+}, Mn^{2+} and Zn^{2+}. Titration of PhTET3 activity against

Table 2. Effect of metal cations on PhTET3 activity.

Metal	Relative activity (%)
None	100
Ca^{2+}	55.3
Cd^{2+}	35.4
Co^{2+}	236.7
Cu^{2+}	12.1
Mg^{2+}	63.0
Mn^{2+}	13.7
Zn^{2+}	15.1

Lys-pNA with Co^{2+} revealed that cobalt optimally enhanced the enzyme activity at 0.1 mM concentration (results not shown); hence, this concentration was routinely used in PhTET3 activity assays.

PhTET3 is a metalloaminopeptidase

In order to further characterize the enzymatic activity, various potential inhibitors were tested. The effects of several chemical agents on PhTET3 activity are summarized in Table 3. The enzyme was clearly inhibited in the presence of amastatin, a typical inhibitor of aminopeptidases, and EDTA, a metal ion-chelating agent. Bestatin, another well-known aminopeptidase inhibitor, decreased the enzymatic activity as well, but its inhibiting effect was clearly lower. Weak inhibition was also detected in the

Table 3. Effect of different chemical agents on PhTET3 activity.

Chemical agent	Concentration (mM)	Relative activity (%)
None	–	100
Amastatin	0.0005	63.6
	0.001	39.0
	0.002	23.4
Antipain	0.2	101.8
Bestatin	0.1	89.4
	0.2	78.7
	0.5	66.4
Chymostatin	0.2	95.3
	0.5	94.7
	1	78.4
E-64	1	97.2
EDTA	0.01	76.2
	0.02	14.7
	0.05	10.6
Leupeptin	0.02	103.5
Pefabloc SC	1	91.6
	2	90.2
	5	80.1
Pepstatin	0.01	98.2
Phosphoramidon	1	99.9
Puromycin	0.5	87.8
	1	90.5

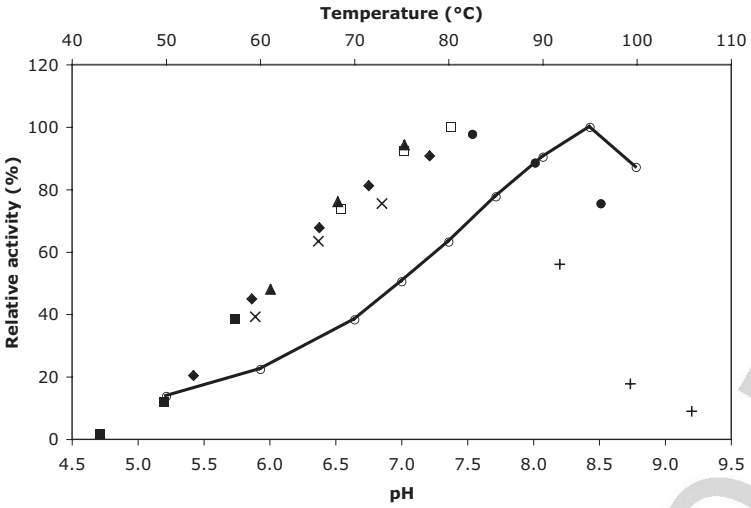


Fig. 4. Effect of pH and temperature on PhTET3 activity. Evolution of PhTET3 activity as a function of pH in (■) MES, (◆) TRIS, (×) PIPES, (▲) HEPES, (□) TAPS, (●) CHES and (+) CAPS buffers, and as a function of temperature (m).

presence of the chymotrypsin inhibitor chymostatin, but only when this agent was used at very high concentrations (its effective inhibitory concentration is normally in the range 10–100 μM), and when the serine proteases inhibitor pefabloc SC was added to the reaction mixture. PhTET3 was insensitive to the papain and trypsin inhibitor antipain, to the cysteine proteases inhibitor E-64, to the serine and cysteine proteases inhibitor leupeptin, to the aspartate proteases inhibitor pepstatin, and to the metalloendopeptidases inhibitor phosphoramidon. Puromycin, an inhibitor of some exopeptidases, had practically no effect on the activity of PhTET3.

PhTET3 is a highly thermostable neutral peptidase

Using the PhTET3 activity assay with Lys-pNA as a substrate, the enzyme showed optimal amidolytic activity at pH 7.4 in TAPS buffer, and more than 50% of the optimum activity was measured between pH 6 and pH 8.5 (Fig. 4). CAPS buffer seems to be an inhibitor of PhTET3, so activity could not be measured more accurately at higher pHs. However, the decreased activity could also be related with the formation of an insoluble complex by the Co^{2+} at basic pH that would, consequently, diminish the availability of the cation. The temperature dependence of PhTET3 activity at pH 7.4 is also shown in Fig. 4. The optimal temperature for enzyme activity was found to be 95°C. At temperatures lower than 60°C, less than 20% of the maximum activity was detected. Thermostability of PhTET3 was studied as well. The enzymatic activity decay obeyed first-order kinetics and displayed half-lives of 27.0 min and 17.83 h when incubated at 100° and 80°C respectively (Table 4).

Thermostability of PhTET3 quaternary structure studied by SANS

The thermostability of the PhTET3 quaternary structure was studied by SANS. Two samples of PhTET3 were incubated for 14 h at 20°C and 90°C before the experiments, which were carried out at the incubation temperature. The obtained curves show high similarity in their overall shape (Fig. 5). Using the Guinier approximation, the radii of gyration were calculated to be 49 ± 1 Å at 20°C and 52 ± 1 Å at 90°C. Given the resemblance of the shape of the curves and the relatively small change in the radius of gyration, we conclude that the incubation at 90°C did not lead to desoligomerization of the PhTET3 dodecamer. In order to verify that PhTET3 forms dodecameric tetrahedral particles in solution, a theoretical scattering curve from the pdb file of PhTET3 was calculated. The experimental and theoretical curves superpose very well (data not shown). Furthermore, the theoretical radius of gyration was evaluated to be 48 Å. Thus, the SANS data confirm that PhTET3 is dodecameric and has a tetrahedral structure in solution.

Discussion

The cytosol of Archaea contains a collection of ATP-independent cytosolic peptidases that cover a broad

Table 4. Effect of different temperatures on half-life of PhTET3.

Temperature (°C)	Half-life
100	27.0 min
95	55.1 min
90	03.44 h
80	17.83 h

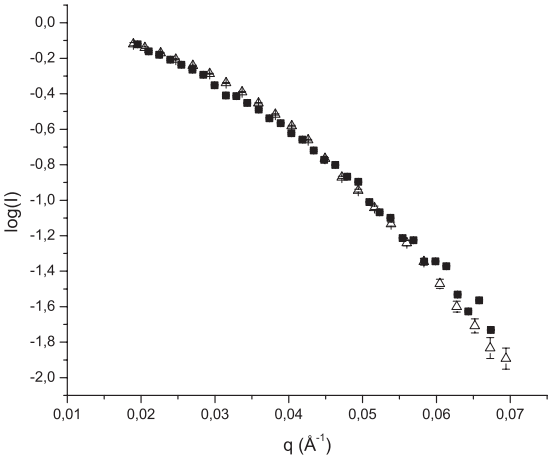


Fig. 5. Logarithm of the SANS intensities of PhTET3 at 20°C (Δ) and 90°C (■) as a function of the scattering vector q ($q = 4\pi\ l^{-1} \sin 2\theta$, λ being the wavelength and θ the scattering angle). Within the experimental errors, both curves show high similarity. Where no error bars can be seen, the error is smaller than the symbol of the curve.

range of activities (Ward *et al.*, 2002; De Castro *et al.*, 2006). The biological functions of these peptidases extend from very specific cleavages (i.e. transit peptide elimination, protein modifications) and energy metabolism (external and internal peptide digestion) to their participation in the protein destruction pathways (elimination of the peptides produced by the proteasome and the other ATP-dependent proteases). The large molecular assemblies capable of manipulating a broad range of polypeptide substrates in an ATP-independent manner represent a small fraction of the cytosolic peptidases and their physiological function is still not clear. In Archaea, two exopeptidase systems have been identified: TRI and TET (Tamura *et al.*, 1996; Franzetti *et al.*, 2002). The two systems exclude themselves in the genomes (except in *Pyrobaculum aerophilum*) and it has been suggested that they fulfil the same biological function (Schoehn *et al.*, 2006).

In *P. horikoshii*, two TET peptidases that self-assemble in dodecameric tetrahedral complexes had been described so far (Russo and Baumann, 2004; Durá *et al.*, 2005; Schoehn *et al.*, 2006). The crystallographic structure of the complex formed by the PH1821 protein (PhTET3) is described in this paper. Like PhTET1 and PhTET2, the structure of PhTET3 reveals a 12-subunit self-compartmentalized particle enclosing 12 catalytic sites that are typical for the M42 metallo-peptidase family (Rawlings *et al.*, 2008). The structural comparison of the *P. horikoshii* TET proteins uncovers a striking similarity between them. These complexes are formed by interactions between dimers to build up a tetrahedral edifice with comparable dimensions. Four large pores, defined by three dimers are situated in each facet of the tetrahedra and give access to the interior of the systems. On each of the four apices, three monomers define a catalytic chamber leading to a small orifice, hidden by aromatic residues in the case of PhTET2 and PhTET3. The characterization of PhTET3's catalytic activity was performed and revealed that, as PhTET1 and PhTET2, PhTET3 displays a strict aminopeptidase behaviour, and that it is also activated by cobalt ions. Taken together, these results show that a third TET molecular machine exists in *P. horikoshii*. The biological signification of such an apparent functional redundancy inside the same peptidase family is intriguing and prompted us to an in-depth investigation.

PhTET1 was reported to be an aminopeptidase with peptide deblocking activity, and it cleaves preferentially acidic residues and Ala in N-terminal position (Ando *et al.*, 1999; Schoehn *et al.*, 2006; M. A. Durá, unpublished). PhTET2 is a fairly unspecific leucyl aminopeptidase (Durá *et al.*, 2005). Within peptidases of the same family, few changes in the active site environment can lead to a different type of activity. For example, family M28 in clan MH contains amino- and also carboxypeptidases. Moreover, despite high sequence identity and/or structural similarity with peptidases, some proteins display a completely different function. In this regard, and like other families in the clan, family M42 includes enzymes that are not peptidases, such as an endoglucanase (Rawlings *et al.*, 2008). These facts triggered us to characterize in detail the enzymatic activity of the purified PhTET3 protein complex and to compare it with the other PhTET peptidases, taking also into account their respective structural features.

PhTET3 activity was first assayed on small chromogenic peptidase substrates and on peptides, and we found that the protein can cleave at least nine different amino acid residues from the N-terminus and shows optimal amidolytic activity against Lys-pNA (Table 1). Therefore, PhTET3 is a quite specific lysyl aminopeptidase. This feature distinguishes clearly PhTET3 from PhTET1 and PhTET2. Interestingly, blocking the substrate N-terminus with an acetyl group (i.e. acetyl-Leu-pNA, a PhTET1 substrate according to Ando *et al.* 1999) prevented hydrolysis by PhTET3, indicating that PhTET3 is not a deblocking aminopeptidase. The absence of activity against N-terminal blocked peptides has also been reported for PhTET2 (Durá *et al.*, 2005) and indicates that both proteins have a strict requirement for a free amino terminus and that they are true exopeptidases devoid of endopeptidase activity. Moreover, PhTET3 acted on Ala-Ala-pNA and Ala-Ala-Ala-pNA cleaving only amino acid residues from the N-terminus; therefore, PhTET3, like PhTET1 and PhTET2 (Durá *et al.*, 2005; Schoehn *et al.*, 2006), seems

to lack di- and tripeptidyl activity and can be classified as a strict aminopeptidase. These results are strengthened by the fact that PhTET3 activity was not affected by the presence of endoprotease inhibitors such as antipain or phosphoramidon, but it was inhibited by the typical aminopeptidase inhibitors amastatin and bestatin (Table 3). The absence of clear puromycin inhibition can be explained by the nature of this compound: it is not a peptide, but a nucleoside, so its structure may not fit easily in the PhTET3 catalytic site, as was described for PhTET2 (Durá *et al*., 2005). Additionally, PhTET3 seems to be unable to act on aminoacyl-proline bonds, as no cleaving of the alanyl residue was detected using Ala-Pro-pNA as a substrate. This lack of iminopeptidase activity is also a feature of PhTET1 (Ando *et al*., 1999) and PhTET2 (Durá *et al*., 2005).

The analysis of the PhTET3 action on Ala-Ala-pNA and Ala-Ala-Ala-pNA also shows that the levels of the products do not increase in parallel and that the ratio among their concentrations changes with time (Fig. 4). Thus, degradation of the longest peptides involves the accumulation of shorter intermediary peptides, although they are substrates (e.g. Ala-pNA, Ala-Ala-pNA). In other words, PhTET3 does not degrade peptides in a processive fashion, as it separates from the polypeptide products after each cleavage. This finding also agrees with previous results on PhTET1 (Ando *et al*., 1999; Schoehn *et al*., 2006) and PhTET2 (Durá *et al*., 2005).

In this paper the optimal conditions for PhTET3 enzymatic activity have been also determined. All three PhTETs are cobalt-activated (Ando *et al*., 1999; Durá *et al*., 2005; Table 2). Interestingly, these results, together with the published enzymatic characterizations of two aminopeptidases, a carboxypeptidase, and a prolidase from *P. furiosus* (Tsunasawa *et al*., 1997; Ghosh *et al*., 1998; Cheng *et al*., 1999; Story *et al*., 2005) suggest that cobalt activation is a hallmark for metallopeptidases from Pyrococcales. This feature may correspond to a higher affinity for cobalt ions in the catalytic sites under high temperature and/or to an adaptation associated to the environment, especially in the case of *P. horikoshii*, as deep sea vents are especially rich in cobalt (Llanos *et al*., 2000). The presence of zinc in the active site of PhTET3 could be an artefact of the recombinant expression in *Escherichia coli*.

In its natural environment, *P. horikoshii* is exposed to high temperatures, and therefore, PhTETs are highly thermostable. PhTET3 displays a half-life of 3.44 h at 90°C (Table 4), somewhat longer than that of PhTET2 (3 h; Durá *et al*. 2005) and comparable to that of PhTET1, which loses around 20% of its activity after incubating at 90°C for 2 h (Ando *et al*., 1999). Also, the optimal temperatures for the enzymatic activity reflect those at which *P. horikoshii* grows better. PhTET3, like PhTET1 (Ando *et al*., 1999), acts optimally at around 95°C (Fig. 4), while PhTET2 shows an optimum at 100°C or maybe higher (Durá *et al*., 2005). However, compared with PhTET1 and PhTET2 (Ando *et al*., 1999; Durá *et al*., 2005), PhTET3 retains higher activity at low temperatures. The structural stability of the PhTET3 edifice in extreme temperature conditions was directly measured by SANS. The data analysis shows that the quaternary structure of PhTET3 remains unaffected even when the complex is incubated at 90°C for long periods (Fig. 5). These results establish, for the first time, that the large PhTET edifices are extremely robust and their assemblies can exist for long periods under extreme physiological conditions in the cytosol of *Pyrococcus*. Further studies on TET complexes originating from non-thermophilic Archaea will be necessary to identify the structural determinants of PhTET3 stability.

As a result of the biochemical study performed on PhTET3 and reported in this paper we can conclude that the three PhTET complexes display complementary affinities towards polypeptidic N-termini and are therefore not redundant in the cytosol. A cautious examination of the internal architecture of the particles reveals the presence of specificity pockets adjacent to the active sites in all three PhTETs and intended to accommodate the side-chain of the N-terminal residue of the substrate. In PhTET3, the negative electrostatic charge of the pocket (Fig. 2A, left), as well as its relatively large size, explains the observed preference of this enzyme for substrates with a large and mainly positively charged amino acid residue at the N-terminus (Lys, Arg, Glu, Leu, see Table 1). In PhTET1 the positively charged specificity pocket (Fig. 2A, middle) can easily accommodate negatively charged side-chains (Glu, Asp). For PhTET2, optimal catalytic activity has been detected for the hydrolysis of substrates with small hydrophobic or neutral amino acid residues in N-terminal position, such as Leu, Met, Ile, Ser, Ala (Durá *et al*., 2005). This can be attributed to its smaller and hydrophobic specificity pocket (Fig. 2A, right).

In a previous paper, we proposed a model to explain the mode of action of the TET molecular machines (Schoehn *et al*., 2006). The results of the structural and biochemical characterization of PhTET3 and its comparison with the other *P. horikoshii* TET proteins strengthen and extend that model. In it, four processing systems would cross in the middle of the particle. Each system would consist of an entry pore situated in the facet of the tetrahedron. This pore would be prolonged by a channel that would lead the polypeptides towards the centre of the particle. There, a series of mobile loops would allow addressing the N-terminus towards one of the four catalytic chambers. The comparative structural analysis of the PhTET structures supports this vectorial model for polypeptide

trafficking. First, it reveals that, in the three particles, the electrostatic charge distribution within the channels is highly contrasted (Fig. 2C) and that most of the charged residue side-chains constituting the channel walls are not implicated in any ionic interaction with neighbouring side-chains. This is consistent with an entry channel aimed to maintain the polypeptide chain (also containing different types of charged areas) in continuous motion towards the interior of the particle. Second, the sequence alignments show that the positive charged lysyl residues that form a central 'spherical' layer in PhTET1, and that are part of the loops that are clearly visible at the centre of this particle (Schoehn *et al.*, 2006; Fig. 2B, middle), are conserved in all three PhTETs (K122, K131/K132 and K133 in PhTET1, 2 and 3 respectively; see Fig. 1B). These loops appear to be highly disordered and cannot be seen in the PhTET2 and PhTET3 structures. This can be related to their high mobility. These loops correspond to residues S119 to A133 in PhTET2 (Borissenko and Groll, 2005), and to residues R128 to K136 in PhTET3, and their absence in the solved structures of these complexes would explain why the PhTET2 and PhTET3 edifices seem to enclose a big central cavity while PhTET1 contains a clear system of four entry channels prolonged by four catalytic chambers (Schoehn *et al.*, 2006). As already described, these catalytic chambers are mainly positively charged in the three PhTETs (Fig. 2B). Once the polypeptide would have been addressed by the mobile loops to one of those chambers, electrostatic repulsion would orient the positively charged N-terminus towards the active sites. These are negatively charged in all three PhTETs due to the conserved active site residues. The next step would be the placement of the N-terminal residue side-chain in the specificity pocket, leaving the substrate peptide in the correct position for hydrolysis.

How the released amino acids would be expelled from the particles is more difficult to infer. In the case of PhTET2, it has been suggested that three tiny pores that are situated around the facet large openings and connected to the catalytic site serve as exits for the free amino acids (Borissenko and Groll, 2005). PhTET1 structural analysis suggested another option, as those three narrow pores are absent. In this case, it was proposed instead that the free amino acids were ejected by the small pores located in the centre of the apices (Schoehn *et al.*, 2006). In the PhTET3 edifice there are no tiny pores in the facets and the apices are virtually closed by the Y227 side-chains. However, these could be displaced by a dynamic process associated with the N-terminal cleavage. This movement would unmask an exit hole whose dimensions would be well adapted for the free amino acid expulsion. This motility of the protein is supported by the non-processive nature of the PhTET enzymes that implies the detachment of the peptide moiety from the active site once the N-terminal residue has been cleaved, possibly to allow the discharge of the free amino acid in the catalytic chamber towards the apex hole. At the same time, the presence of a closed exit hole when the enzyme is not expelling the free amino acid products would avoid the entry into the particle of small peptides difficult to orient or not destined to destruction that could block or diminish the catalytic efficiency of the complex.

In conclusion, the data presented in this paper show that, in the cytosol of *P. horikoshii*, three TET-like molecular machines combine their different substrate specificities to form an integrated, thermostable polypeptide destruction system. Such a cooperative action between cytosolic peptidase activities had only been demonstrated for the three TRI-interacting aminopeptidases that combine their specificities to break down the small peptides released by the TRI complex (Tamura *et al.*, 1998). Its broad activity towards polypeptides indicates that the archaeal TET system is likely to be a key player in many cellular functions that require polypeptide breakdown, whatever the origin of the peptides. The cellular role of the TET particles could therefore extend from protein quality or protein turnover control, by working beyond the 20S proteasome, to energy metabolism by digesting the peptides imported from the external environment. In this respect, it is possible that the TET system is particularly important for the heterotrophic lifestyle of the sea microbes that rely on external peptides as a nutrient source (Karl, 2002). This would explain the presence of different TET homologous proteins, with probable optimized complementary specificities, in the genomes of all known Pyrococcales.

Experimental procedures

Protein expression and purification

We used the *E. coli* strain BL21-CodonPlus (DE3)-RIL (Stratagene) for expression of recombinant wild-type PhTET3. Two clone colonies were inoculated in 50 ml of LB medium [20 g of LB broth (Sigma) l^{-1} of deionized water], containing ampicillin (100 mg l^{-1}) and chloramphenicol (34 mg l^{-1}). The cells were grown with shaking at 37°C overnight. Ten millilitres of this subculture was added to 1 l of LB medium supplemented with 100 mg of ampicillin. After incubation with shaking at 37°C until the A_{600} reached 0.6–1.0, the induction was carried out by adding isopropyl β-D-thiogalactopyranoside at a final concentration of 0.05 mM and shaking for 4 h at 37°C. The induced cells were harvested by centrifugation. The pellets were suspended in 50 ml of 50 mM Tris-HCl, 150 mM NaCl, 0.1% Triton X-100, pH 8.0 and maintained at –80°C until used. Cell suspensions were thawed at room temperature and added with 12.5 mg of lysozyme (Euromedex), 2.5 mg of DNase I grade II (Roche), 10 mg of RNase (Roche), 50 mg of Pefabloc SC (Roche) and 0.5 ml of 2 M $MgSO_4$. The disruption of the cells was achieved by sonication in a Branson sonifier 150 at 4°C. Ten 30 s bursts at intensity 10 with

intermediary pauses of 30 s were employed. The crude extract was heated at 85°C for 15 min and then it was clarified by centrifugation at 17000 *g* for 1 h at 4°C. Supernatant NaCl and Tris concentrations were adjusted to 0.2 M and 20 mM respectively, the pH was adjusted to 7.5, and the resulting extract was injected in a DEAE Sepharose CL-6B (15 cm × 1.6 cm) column (Pharmacia Biotech), previously equilibrated with 20 mM Tris-HCl, 0.2 M NaCl, pH 7.5. The protein was eluted at 2 ml min^{-1} with a linear NaCl gradient from 0.2 to 0.45 M in 200 ml. The chromatographic separation was done in an ÄKTAprime system (Amersham Biosciences). Fractions of 8 ml were collected and those with Lys-pNA hydrolytic activity were pooled. The pool NaCl concentration was adjusted to 0.3 M before being applied in two batches to a Mono Q HR 5/5 column (Amersham Pharmacia Biotech) previously equilibrated with 20 mM Tris-HCl, 0.3 M NaCl, pH 7.5. The column was run at a flow rate of 1 ml min^{-1} in an ÄKTApurifier system (Amersham Biosciences) and protein elution was achieved by a linear NaCl gradient from 0.3 to 0.65 M in 20 ml. Fractions of 0.5 ml were collected and those showing lysyl aminopeptidase activity were pooled and concentrated up to 1 ml by centrifugation in an Amicon Ultra Ultracel-30k membrane filter (Millipore). Concentrated fractions were loaded onto a HiPrep 16/60 Sephacryl S-300 HR column (Amersham Biosciences) previously equilibrated with 20 mM Tris-HCl, 0.15 M NaCl, pH 7.5. The column was run at a flow rate of 0.5 ml min^{-1} in an ÄKTAprime system (Amersham Biosciences) and 1 ml fractions were collected. Fractions containing Lys-pNA hydrolytic activity, and that were not contaminated according to the chromatogram recorded following the absorbance at 280 nm, were pooled and kept at 4°C after concentration to 7 mg ml^{-1}.

Determination of protein concentration

Protein concentration was measured using the Bio-Rad protein assay reagent (Bio-Rad) and bovine serum albumin as standard. A correction factor was applied to pure PhTET3 protein samples. This factor was calculated after determining the protein concentration of pure PhTET3 samples by quantitative amino acid analysis. For that, the samples were dried and hydrolysed at 110°C in constant-boiling HCl containing 1% (v/v) phenol, for 24 h under reduced pressure and in the absence of oxygen. Amino acids were analysed on a model Biochrom 30 amino acid analyser, with the standard sodium citrate eluting buffer system. Calibrations were made with standard solutions of all the amino acid except tryptophan. Therefore, all the values reported in this paper refer to real PhTET3 protein concentrations.

Protein crystallization

To crystallize PhTET3 EasyXtalTool X-Seal plates (Qiagen) were used and crystallization was achieved by the hanging drop method. PhTET3 was found to crystallize with a mother liquor containing 85 mM Tris-HCl, pH 8.5, 21.25% polyethylene glycol 3350 and 170 mM ammonium sulphate. For crystallization, 1 ml of mother liquor was placed in the well of the crystallization plate and the drop was formed by mixing 2 µl of protein solution at 7 mg ml^{-1}, 2 µl of mother liquor and 0.5 µl of a 40% solution of γ-butyrolactone.

Table 5. Data collection and refinement statistics.

Diffraction data	
Space group, cell parameters	I23, 132.241 Å, 90°
Resolution limits (Å)	33.0–1.9 (2.02–1.90)
Reflections measured	433304 (61427)
Unique reflections	30379 (5057)
Completeness	99.9% (99.7%)
R_{merge}	6.8% (42.6%)
I/σ(I)	22 (6)
Data redundancy	14.3 (12.1)
Refinement	
R_f	18.1% (22.6%)
R_{free}	19.3% (25.0%)
Rmsd bond length/angles	0.005 Å/1.4°
Φ/Ψ plot	
Most favoured	90.2% (258)
Additionally allowed	9.4% (27)
Generously allowed	0.0% (0)
Disallowed	0.3% (1)
Estimated co-ordinate error from cross validated Luzzati plot	0.21 Å

Diffraction data collection and structure determination

X-ray diffraction data were collected on the ID 23-1 beamline at the European Synchrotron Radiation Facility on a single crystal of PhTET3. For cryoprotection, the crystal was allowed to soak for 10 s in a solution containing the mother liquor and 15% glycerol. Data frames were processed with XDS (Kabsch, 1988) and the 3D structure of PhTET3 was solved by molecular replacement (Rossmann, 1972) using the 3D structure of the monomer of PhTET2 (PDB id 1Y0Y) that has a sequence identity of 48% with PhTET3. The calculations were performed with PHASER (Read, 2001) using all available diffraction data. The molecular replacement solution has a *Z*-value of 25.1, a log-likelihood gain of 437 and a proper packing into the I23 cell, which reflects the arrangement of the monomers into homo-dodecameric particles. Statistics of data collection and refinement are available in Table 5. Model refinement was carried out with CNS (Brünger *et al.*, 1998) using energy minimization, individual isotropic temperature factor adjustment and annealing in torsion-angle space. Model rebuilding sessions were done using COOT (Emsley and Cowtan, 2004). The quality and stereochemistry of the model were checked with CNS (Brünger *et al.*, 1998) and PROCHECK (Laskowski *et al.*, 1993). APBS (Baker *et al.*, 2001) and pdb2pqr (Dolinsky *et al.*, 2004) were used to perform electrostatic calculations. All figures were prepared using PyMOL (DeLano, 2002).

The atomic co-ordinates and measured structure factor amplitudes for PhTET3 have been deposited in the Protein Data Bank with accession code 2VPU.

Determination of PhTET3 activity on synthetic chromogenic and fluorogenic compounds

PhTET3 hydrolytic activity on synthetic chromogenic and fluorogenic compounds was determined using different aminoacyl-pNAs, aminoacyl-AMCs and peptides (listed in Table 1) through the following standard assay procedure.

Reactions were initiated by addition of the enzyme (0.1 μg ml^{-1}) to a pre-warmed mixture containing the chromogenic or fluorogenic compound in 50 mM TAPS, 150 mM KCl, 0.1 mM $CoCl_2$, pH 7.4 at 85°C and covered by a layer of mineral oil to avoid water evaporation. To circumvent the problem of insoluble blue complex formation by the Co(II) ions at basic pH, the $CoCl_2$ was added when the buffer was at working temperature. Incubations were performed at 85°C for 0.5, 5 or 30 min and reactions were stopped by adding an equal volume of 0.1 M acetic acid and cooling to 0°C. The quantity of pNA, AMC or β-naphthylamine liberated was measured in a VICTOR 1420 Multilabel Counter (Wallac). For pNA, absorbance was determined at 405 nm, and for AMC and β-naphthylamine, fluorescence was measured using excitation and emission wavelengths of 355 and 460 nm respectively. Due to the higher signal response of the AMC derivatives with respect to the pNA ones, the former were used at lower concentrations and incubation times were accordingly reduced. Four replicates and four enzyme blanks were assayed for each experimental point.

For the study of Ala-Ala-pNA and Ala-Ala-Ala-pNA degradation by PhTET3, the enzyme (0.1 μg ml^{-1}) was incubated with the substrate (5 mM) using the standard assay conditions. At determined times, 60 μl aliquots were removed from the reaction mixture and added to 150 μl of cold acetonitrile in order to stop the reaction and precipitate the protein, which was removed by centrifugation. A 196 μl aliquot of supernatant was dried under vacuum. Solids were then dissolved in 140 μl of 0.065% (v/v) trifluoroacetic acid, 2% (v/v) acetonitrile in water and the sample was analysed by reverse phase chromatography as described by Schoehn *et al.* (2006). Each sample was accompanied by an appropriate enzyme blank.

Effect of metal cations on PhTET3 activity

The metal cations (listed in Table 2) were incubated at 1 mM concentration with the enzyme (0.15 μg ml^{-1}) and the substrate (5 mM Lys-pNA) at 85°C in 50 mM TAPS, 150 mM KCl, pH 6.7 at 85°C, for 5 min. The reactions were stopped by dilution with the same volume of 0.1 M acetic acid and cooling to 0°C, and pNA absorbance was measured as previously described. Four replicates and four enzyme blanks were assayed for each experimental point.

Effect of temperature and pH on PhTET3 activity

The effect of pH on PhTET3 activity was determined in the pH range from 4.7 to 9.2. The buffers used were: MES, pH 4.7–5.7; PIPES, pH 5.9–6.9; HEPES, pH 6.0–7.0; TRIS, pH 5.4–7.2; TAPS, pH 6.5–7.4; CHES, pH 7.5–8.5; and CAPS, pH 8.2–9.2. All buffers were used at 50 mM concentration and the pH was adjusted with NaOH in all cases except with TRIS, where HCl was employed. The given pH values are at 85°C. The effect of temperature on the activity was measured in the range from 50°C to 99.9°C. In all cases, the enzyme (0.1 μg ml^{-1}) was incubated with 5 mM substrate (Lys-pNA). To assess the effect of pH on activity, incubations were done at 85°C for 5 min in the indicated buffers to which 150 mM KCl and 0.1 mM $CoCl_2$ were added. When the effect of temperature was determined, incubations were performed in 50 mM TAPS, 150 mM KCl, 0.1 mM $CoCl_2$, pH 7.4 at the working temperature. The reactions were stopped by adding the same volume of 0.1 M acetic acid and cooling to 0°C, and pNA absorbance was measured. Four replicates and four enzyme blanks were assayed for each experimental point.

Thermal stability of PhTET3 activity

The effect of temperature on PhTET3 stability was assayed by incubating the enzyme (2 μg ml^{-1}) at several temperatures (listed in Table 4) in 50 mM TAPS, 150 mM KCl, pH 7.4 at the working temperature. Aliquots were taken at different time intervals and the remaining aminopeptidase activity was measured following the standard assay procedure. Four replicates (samples + enzyme blanks) were assayed for each experimental point. The remaining activity at each time interval was expressed as a percentage of the activity present at zero time (100%). Half-lives were calculated from a first-order exponential decay fit to the six experimental points obtained for each temperature.

Effect of chemical agents on PhTET3 activity

The enzyme (0.1 μg ml^{-1}) was incubated together with the chemical agents listed in Tables 3 and 5 mM Lys-pNA following the standard assay procedure. pNA absorbance was measured as described. Four replicates and four enzyme blanks were assayed for each experimental point.

SANS experiments

Two samples of PhTET3 were studied by SANS. Both samples were measured in 20 mM Tris-HCl buffer adjusted to pH 7.5 at the temperature of the experiment. One sample with 2.3 mg of PhTET3 ml^{-1} was incubated at 90°C overnight before the experiment and additionally contained 150 mM KCl. The other sample contained 3.1 mg of PhTET3 ml^{-1}. The experiments were carried out on the small-angle diffractometer D22 at the Institut Laue-Langevin (ILL), Grenoble, France. For the experiments, 160 μl of sample was pipetted into Hellma QS quartz cuvettes with an optical path length of 1 mm and placed in a multisample holder whose temperature was maintained constant throughout the experiment (90°C or 20°C). For the experiment at 20°C, an instrument configuration of detector/collimator = 2/2 m was chosen and the sample was measured for 90 min. For the experiment at 90°C the configuration detector/collimator was 3/4 m and the exposure time was 60 min. The wavelength was 7 Å in both cases. In addition to the samples, the corresponding buffers, a boron B_4C standard, an empty quartz cuvette and the neutron beam were measured. The sample and buffer raw intensities were normalized to the incoming neutron flux and corrected for detector efficiency (buffer reference), electronic background (boron standard) and sample holder scattering (empty quartz cuvette) using a program suite developed at the ILL (Gosh *et al.*, 2006). Subsequently, the scattered intensities were summed up azimuthally. Finally, the background intensities were subtracted from the respective sample intensities with the program PRIMUS (Konarev *et al.*, 2003). The data were

exploited in a *q*-range from 0.02 Å^{-1} to 0.07 Å^{-1} for the sample at 20°C and from 0.016 Å^{-1} to 0.07 Å^{-1} for the sample at 90°C. The radii of gyration were calculated using the Guinier approximation (Guinier, 1939) in a *q*-range from 0.019 Å^{-1} to 0.032 Å^{-1} and from 0.016 Å^{-1} to 0.028 Å^{-1} for the samples at 20°C and 90°C respectively. Theoretical scattering curves were calculated using the program CRYSON (Svergun *et al.*, 1998).

Acknowledgements

This work was funded by grants from the French Marine Genomics Network, the CNRS 'Origin of Planets and Life' interdisciplinary program and the French National Research Agency (BLAN07-3_204002). M.A.D. was supported by a Marie Curie Intra-European Fellowship (European Community Sixth Framework Programme) and E.R. by a PhD scholarship from the French Ministry for Research and Technology. We thank Drs Alain Roussel and Véronique Réceveur-Bréchot for kindly providing us with the plasmid encoding PhTET3, the European Synchrotron Radiation Facility staff for their assistance in X-ray data collection, the ILL staff and Dr Giuseppe Zaccaï for their help in the SANS experiments, and Jean-Pierre Andrieu for performing the amino acid analysis.

References

Ando, S., Ishikawa, K., Ishida, H., Kawarabayasi, Y., Kikuchi, H., and Kosugi, Y. (1999) Thermostable aminopeptidase from *Pyrococcus horikoshii*. *FEBS Lett* **447:** 25–28.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98:** 10037–10041.

Baumeister, W., Walz, J., Zühl, F., and Seemüller, E. (1998) The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92:** 367–380.

Borissenko, L., and Groll, M. (2005) Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J Mol Biol* **346:** 1207–1219.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54:** 905–921.

Cheng, T.C., Ramakrishnan, V., and Chan, S.I. (1999) Purification and characterization of a cobalt-activated carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Protein Sci* **8:** 2474–2486.

Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.M., Moras, D., and Tarnus, C. (1994) Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure* **2:** 283–291.

De Castro, R.E., Maupin-Furlow, J.A., Gimenez, M.I., Herrera Seitz, M.K., and Sanchez, J.J. (2006) Haloarchaeal proteases and proteolytic systems. *FEMS Microbiol Rev* **30:** 17–35.

DeLano, W.L. (2002) *The PyMOL Molecular Graphics System*. San Carlos, CA: DeLano Scientific. [WWW document]. URL http://www.pymol.org

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32:** W665–W667.

Durá, M.A., Receveur-Brechot, V., Andrieu, J.-P., Ebel, C., Schoehn, G., Roussel, •••, and Franzetti, B. (2005) Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Biochemistry* **44:** 3477–3486.

Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132.

Franzetti, B., Schoehn, G., Hernandez, J.-F., Jaquinod, M., Ruigrok, R.W.H., and Zaccai, G. (2002) Tetrahedral aminopeptidase: a novel large protease complex from Archaea. *EMBO J* **21:** 2132–2138.

Ghosh, M., Grunden, A.M., Dunn, D.M., Weiss, R., and Adams, M.W. (1998) Characterization of native and recombinant forms of an unusual cobalt-dependent proline dipeptidase (prolidase) from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* **180:** 4781–4789.

Godfroy, A., Raven, N.D., and Sharp, R.J. (2000) Physiology and continuous culture of the hyperthermophilic deep-sea vent archaeon *Pyrococcus abyssi* ST549. *FEMS Microbiol Lett* **186:** 127–132.

Gonzales, T., and Robert-Baudouy, J. (1996) Bacterial aminopeptidases: properties and functions. *FEMS Microbiol Rev* **18:** 319–344.

González, J.M., Masuchi, Y., Robb, F.T., Ammerman, J.W., Maeder, D.L., Yanagibayashi, M., *et al.* (1998) *Pyrococcus horikoshii* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles* **2:** 123–130.

Gosh, R.E., Egelhaaf, S.U., and Rennie, A.R. (2006) *A Computing Guide for Small-Angle Scattering Experiments*. ••••: ILL06GH05T: Institut Laue Langevin Internal Publications.

Gottesman, S. (2003) Proteolysis in bacterial regulatory circuits. *Annu Rev Cell Dev Biol* **19:** 565–587.

Guinier, A. (1939) La diffraction des rayons X aux très faibles angles: applications à l'étude des phénomènes ultramicroscopiques. *Ann Phys (Paris)* **12:** 161–236.

Kabsch, W. (1988) Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J Appl Crystallogr* **21:** 916–924.

Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22:** 2577–2637.

Karl, D.M. (2002) Nutrient dynamics in the deep blue sea. *Trends Microbiol* **10:** 410–418.

Kirschner, M. (1999) Intracellular proteolysis. *Trends Cell Biol* **9:** M42–M45.

Kisselev, A.F., Akopian, T.N., and Goldberg, A.L. (1998) Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes. *J Biol Chem* **273:** 1982–1989.

Kisselev, A.F., Akopian, T.N., Woo, K.M., and Goldberg, A.L. (1999) The sizes of peptides generated from protein by mammalian 26 and 20S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J Biol Chem* **274:** 3363–3371.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36:** 1277–1282.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26:** 283–291.

Llanos, J., Capasso, C., Parisi, E., Prieur, D., and Jeanthon, C. (2000) Susceptibility to heavy metals and cadmium accumulation in aerobic and anaerobic thermophilic microorganisms isolated from deep-sea hydrothermal vents. *Curr Microbiol* **41:** 201–205.

Lupas, A., Flanagan, J.M., Tamura, T., and Baumeister, W. (1997) Self-compartmentalizing proteases. *Trends Biochem Sci* **22:** 399–404.

Rawlings, N.D., Morton, F.R., Kok, C.Y., Kong, J., and Barrett, A.J. (2008) MEROPS: the peptidase database. *Nucleic Acids Res* **36:** D320–D325. [WWW document]. URL http://merops.sanger.ac.uk/index.htm.

Read, R.J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr D Biol Crystallogr* **57:** 1373–1382.

Rossmann, M.G. (1972) *The Molecular Replacement Method.* New York: Gordon and Breach Science Publisher.

Russo, S., and Baumann, U. (2004) Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J Biol Chem* **279:** 51275–51281.

Saric, T., Graef, C.I., and Goldberg, A.L. (2004) Pathway for degradation of peptides generated by proteasomes: a key role for thimet oligopeptidase and other metallopeptidases. *J Biol Chem* **279:** 46723–46732.

Sauer, R.T., Bolon, D.N., Burton, B.M., Burton, R.E., Flynn, J.M., Grant, R.A., *et al.* (2004) Sculpting the proteome with AAA (+) proteases and disassembly machines. *Cell* **119:** 9–18.

Schoehn, G., Vellieux, F.M.D., Durá, M.A., Receveur-Bréchot, V., Fabry, C.M.S., Ruigrok, R.W.H., *et al.* (2006) An archaeal peptidase assembles into two different quaternary structures: a tetrahedron and a giant octahedron. *J Biol Chem* **281:** 36327–36337.

Snowden, L.J., Blumentals, I.I., and Kelly, R.M. (1992) Regulation of proteolytic activity in the hyperthermophile *Pyrococcus furiosus*. *Appl Environ Microbiol* **58:** 1134–1141.

Stamper, C., Bennett, B., Edwards, T., Holz, R.C., Ringe, D., and Petsko, G. (2001) Spectroscopic and X-ray crystallographic characterization of bestatin bound to the aminopeptidase from *Aeromonas (Vibrio) proteolytica*. *Biochemistry* **40:** 7035–7046.

Story, S.V., Shah, C., Jenney, F.E., and Adams, M.W.W. (2005) Characterization of a novel zinc-containing, lysine-specific aminopeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bact* **187:** 2077–2083.

Svergun, D., Richard, S., Koch, M.H.J., Sayers, Z., Kuprin, S., and Zaccai, G. (1998) Protein hydration in solution: experimental observation by X-ray and neutron scattering. *Proc Natl Acad Sci USA* **95:** 2267–2272.

Tamura, N., Lottspeich, F., Baumeister, W., and Tamura, T. (1998) The role of Tricorn protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell* **95:** 637–648.

Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F., and Baumeister, W. (1996) Tricorn protease – the core of a modular proteolytic system. *Science* **274:** 1385–1389.

Tomkinson, B. (1999) Tripeptidyl peptidases: enzymes that count. *Trends Biochem Sci* **24:** 355–359.

Tsunasawa, S., Izu, Y., Miyagi, M., and Kato, I. (1997) Methionine aminopeptidase from the hyperthermophilic Archaeon *Pyrococcus furiosus*: molecular cloning and overexpression in *Escherichia coli* of the gene, and characteristics of the enzyme. *J Biochem* **122:** 843–850.

Ward, D.E., Shockley, K.R., Chang, L.S., Levy, R.D., Michel, J.K., Conners, S.B., and Kelly, R.M. (2002) Proteolysis in hyperthermophilic microorganisms. *Archaea* **1:** 63–74.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Chapter 4

Oligomerization of PhTET3

4.1 Introduction

In nature oligomeric assemblies are very common. In 2004 only one third of the 452 human enzymes for which the subunit composition was listed in the Brenda database [126] were monomers [127]. There must therefore be a strong advantage in using oligomeric complexes in cellular mechanisms. It has been suggested that large proteins are more stable against denaturation because of a large number of internal interactions, balancing the entropic factors of a reduced number of conformations and solvation. Additionally a smaller surface area will reduce the amount of solvent needed for hydration [128]. Furthermore oligomerization can permit to control enzymatic activity by regulation of the oligomeric state if only one form is active [127, 129]. It has been reported that the ATP-dependent protease Lon from *Mycobacterium smegmatis* shows an equilibrium between a hexamer and lower oligomeric states and that only the hexamer is active [130].
In the genomes of halophilic archaea there is only one gene encoding a TET protease whereas in *P. horikoshii* three TET proteases have been identified. Studies on *Halobacterium salinarum* have shown that TET protein is present in a dimeric form and in a dodecameric form *in vivo* (unpublished results). Depending on the physiological state of the cell, dimers can account for up to 80% of the total TET protein in the cell (unpublished results). When recombinant PhTET3 is purified, it is found exclusively in its dodecameric form. Considering the results of the *in vivo* experiments on *H. salinarum*, it seems likely that PhTET3 exists in different oligomeric states *in vivo* as well. However as the role of the TET proteases in the cell is unknown, it is not clear why a dimeric form exists.
In the case of PhTET3 it seems possible that the subsistence of different oligo-

meric forms might be a way to control enzymatic activity as well. By looking at the crystallographic structure of the enzyme one can see that at a distance of about 10Å from the active sites residues from another subunit are located. These residues are not part of the specificity pocket itself, but form part of the immediate surroundings of the pocket. It seems therefore possible that upon deoligomerization the conformation of the active site residues changes. Furthermore access to active sites of proteases with relatively broad substrate specificity like the TET protease, needs to be regulated inside the cytoplasm to avoid unspecific polypeptide degradation (see section 1.3.1). Hence it is likely that the TET dimers are not active, supporting the hypothesis of oligomerization as a means to control TET activity *in vivo*.
As a first step towards understanding the mechanism of TET oligomerization we have studied the parameters controlling PhTET3 oligomerization *in vitro*.

4.2 Materials and Methods

4.2.1 Production of recombinant PhTET3

The protein was obtained as described in chapter 3.

4.2.2 Analytical Ultracentrifugation

Sedimentation velocity experiments were performed at 42000 rpm and 20 °C on a XL-I analytical centrifuge (Beckman) at the Institut de Biologie Structurale, Grenoble, France. Protein concentration was 0.4 mg/ mL in all samples. Experiments were carried out either with an An-Ti50 (Beckman) 8-hole rotor or with an An-Ti60 (Beckman) 4-hole rotor. Two-channel centerpieces were used with an optical path of 12 mm. All experiments were done with sapphire windows. Scans were recorded at 280 nm with radial spacing of 0.005 cm. The program Sednterp [131] was used to estimate the partial specific volume $\overline{v}$ from amino acid composition as well as the density ρ and viscosity η of the buffer. Data was then analyzed with the program Sedfit [125] using a continuous c(s) distribution model.

4.2.3 Small Angle Neutron Scattering

SANS experiments were carried out on the small angle diffractometer D22 at the Institut Laue Langevin, Grenoble, France. For the experiments 160 µL of sample were pipetted into Hellma QS quartz cuvettes with an optical

path length of 1mm and placed in a multi-sample holder whose temperature was maintained constant throughout the experiment. For the experiment three different instrument configurations were used. Either a detector/collimator=2/2 m, 3/4 m or 4/4 m configuration was chosen. The wavelength was 7Å in all cases. In addition to the samples the corresponding buffers, a boron carbite standard, an empty quartz cuvette and the empty neutron beam were measured. The sample and buffer raw intensities were radially averaged, corrected for electronic background (direct beam interrupted by a boron absorber), normalised to the incoming neutron flux and corrected for detector efficiency by using 1.00mm of H2O as a standard, subtracted of sample holder (quartz cell) scattering corrected for transmission, by using the ILL suite of programs [132]. Finally, the background intensities were subtracted from the respective sample intensities with the program PRIMUS [133].

4.3 Results

4.3.1 Analytical Ultracentrifugation

Calculation of theoretical sedimentation coefficients. All analytical ultracentrifugation (AUC) experiments were carried out at 20 °C using a protein concentration of 0.4 mg/ mL. 20 mM Tris at pH 7.5 was used as a standard buffer to which other solutes were added, except when testing the influence of pH, where the buffers listed in table 4.2 were used. Theoretical sedimentation coefficients for different hypothetical oligomers of PhTET3 were calculated using the equations described in section 2.5.2. The results are resumed in table 4.1.

PhTET3 oligomerization is salt-independent. AUC experiments on PhTET3 samples containing 0M, 0.15M and 0.5M KCl were carried out. The results are shown in figure 4.1. All three samples contain a single sedimenting species that has a sedimentation coefficient of about 16.3S. According to the results listed in table 4.1 this species is identified as dodecameric PhTET3. This shows that the concentration of KCl does not have a significant influence on the oligomeric state of the protein.

Metal-dependant oligomerization. We then tested the influence of EDTA on PhTET3 oligomerization. EDTA is a metal-chelating agent that acts on divalent cations. Before analysis by AUC, the protein was dialyzed eight

	$f/f_0 = 1.20$		$f/f_0 = 1.33$	
SU	$s_{20,w}(S)$	R_H (Å)	$s_{20,w}(S)$	R_H (Å)
1	3.3	27.1	3.0	30.0
2	5.2	34.1	4.7	37.8
3	6.8	39.0	6.2	43.3
4	8.3	43.0	7.5	47.6
6	10.9	49.2	9.8	54.5
12	17.2	62.0	15.5	68.7

Table 4.1: Sedimentation coefficients $s_{20,w}$ and hydrodynamic radius R_H of hypothetical oligomers (with number of subunits SU) of PhTET3 calculated for two different frictional ratios in 20 mM Tris with ρ=0.998 g/c m^3, η=0.0102cP and $\overline{v}$=0.742 ml/g.

days against a buffer containing 20 mM EDTA. A second sample was dialyzed against 20 mM EDTA and 0.5M KCl in order to test whether salt would influence the oligomeric state. In both samples two different oligomers were detected, sedimenting at 4.7S and 7.5S (figure 4.2a). According to the theoretical values (see table 4.1), the species were identified as dimers and tetramers. A large majority of the total protein is in dimeric form (about 82%) with the tetramer accounting for about 13% of the total protein concentration.
We then dialyzed a sample containing PhTET3 dimers and tetramers against Tris buffer for four days. The sample was subsequently analyzed by AUC. Figure 4.2b shows the sample before and after the removal of EDTA. Without EDTA a peak at 15.8S appears, indicating that dodecameric protein has reformed, representing 28% of all protein in solution. However the dimers still account for the majority of the protein mass. Additionally the amount of tetramer present in the solution has decreased to 8%.
After removal of the EDTA we tried to further reoligomerize the sample by dialyzing against solutions containing zinc or cobalt cations with or without NaCl. The results of AUC studies on these samples are shown in figure 4.3. Dialysis against a buffer containing 2 mM $ZnCl_2$ with no salt led to protein aggregation. In the other three samples more dodecamer was generated compared to the sample to which no metal ions were added. The overall composition of the samples was similar in all three conditions with about 43% dimer, 7% tetramer and 42% dodecamer. Adding NaCl did prevent aggregation when used with the zinc cations but did not promote reoligomerization in the sample containing cobalt.

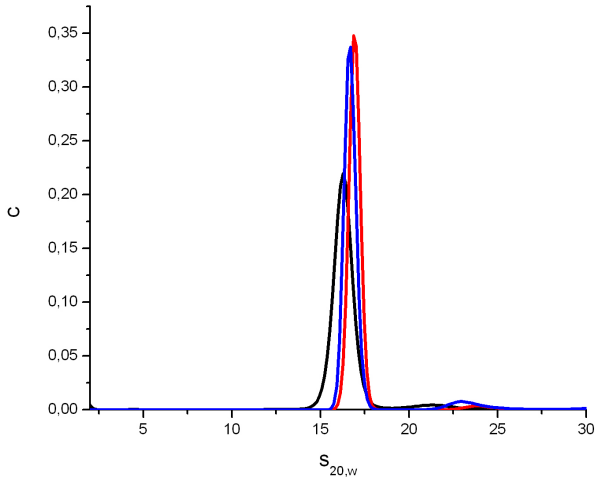
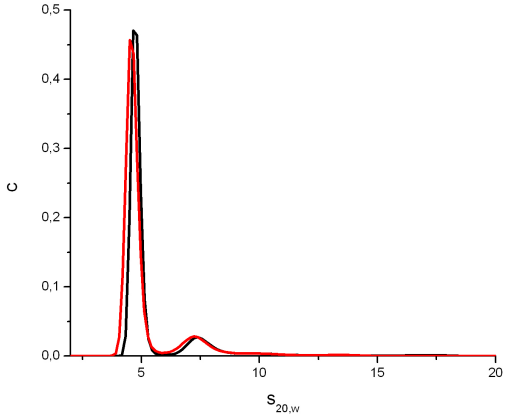


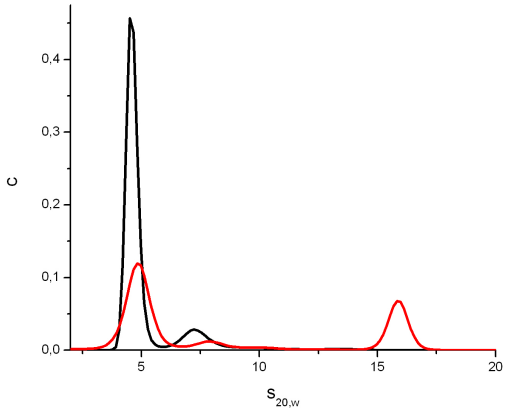
Figure 4.1: Protein concentration as a function of sedimentation coefficient of PhTET3 in three different concentrations of KCl (0M KCl - black, 0.15M KCl - red, 0.5M KCl - blue). All three samples sediment around S=16 $\pm$ 0.3, showing that the protein is dodecameric.

pH-dependent oligomerization. In addition we tested the influence of the pH of the solution on PhTET3 oligomerization. For the experiment the protein was dialyzed overnight against a solution containing the buffer at 50mM concentration and 150mM KCl. Table 4.2 shows the buffer used for each pH. Figure 4.4 shows the percentage of protein that was in dimeric and dodecameric form at different pHs. Whereas the dodecamer of PhTET3 resists acidic pH, it starts to dissociate in basic conditions. However a lower level of dissociation is achieved compared to the experiments using EDTA.

Thermal stability of the PhTET3 dimer. In order to assess the thermal stability of the dimeric form of PhTET3 a sample that was dissociated into dimers by EDTA was heated at 90 ℃ for two hours. EDTA was then removed by dialysis against Tris buffer for four days and the sample was analyzed by AUC. As a control a second dimeric sample was also dialyzed against Tris buffer but without prior heating. As a further control a dodecameric sample was also heated at 90 ℃ for two hours. The dimeric sample that was heated for two hours showed a new species with a sedimentation coefficient of 3.3S. This species which accounts for 17% of the total protein mass in the sample could be monomers according to the theoretical sedimentation coefficients presented in table 4.1. It is also possible that heating



(a) PhTET3 after dialysis against 20mM EDTA, without salt (red line) and with 0.5M KCl (black line).



(b) PhTET3 before (black line) and after (red line) the EDTA was removed by dialyzing against Tris buffer.

Figure 4.2: AUC experiments on the metal-dependant oligomerization of PhTET3. In (a) only dimeric and tetrameric protein is detected after PhTET3 has been dialyzed against 20mM EDTA. After removal of the EDTA by dialysis some dodecamer is reformed (b).

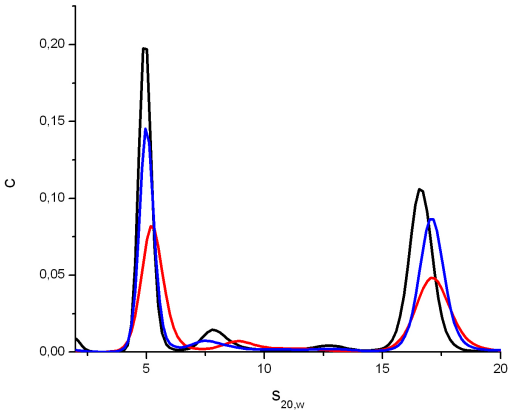


Figure 4.3: PhTET3 after the EDTA was removed by dialysis against Tris buffer with divalent cations. In blue: with 2mM $CoCl_2$; in black: with 2mM $CoCl_2$ and 0.5M NaCl; in red: with 2mM $ZnCl_2$ and 0.5M NaCl. More reoligomerization occurs when divalent cations have been added. However about half of the total protein rests in the dimeric form.

the sample leads to the partial unfolding of dimers, which would lead to a lower sedimentation coefficient. Other species identified are dimers (70%), tetramers (6%) and dodecamers (7%). In the first control sample a majority of the protein is present as dimers (58%). A large population of tetrameric protein is found (18%) and some dodecamer is reformed (18%). The second control sample contains exclusively dodecameric protein. This experiment shows that the dimeric form is less thermostable than the dodecameric form, which is not altered at all by two hours of incubation at 90 °C. Heating the dimer leads to monomer formation or partial unfolding of dimers and reduces reoligomerization into dodecamers upon the removal of EDTA.

4.3.2 Small Angle Neutron Scattering

Calculation of theoretical scattering intensities. To further study PhTET3 oligomerization and in order to gain some structural knowledge on the dimer, we carried out several SANS experiments. Using the x-ray crystallographic structure of PhTET3 theoretical scattering curves for the dodecamer and for the dimer were calculated using the program CRYSON [134]. Owing to the lack of structural data on the dimer, the calculations were performed on the dimer as it is in the dodecameric structure. The calculated curves are represented in figure 4.6 together with experimental data on a sample of

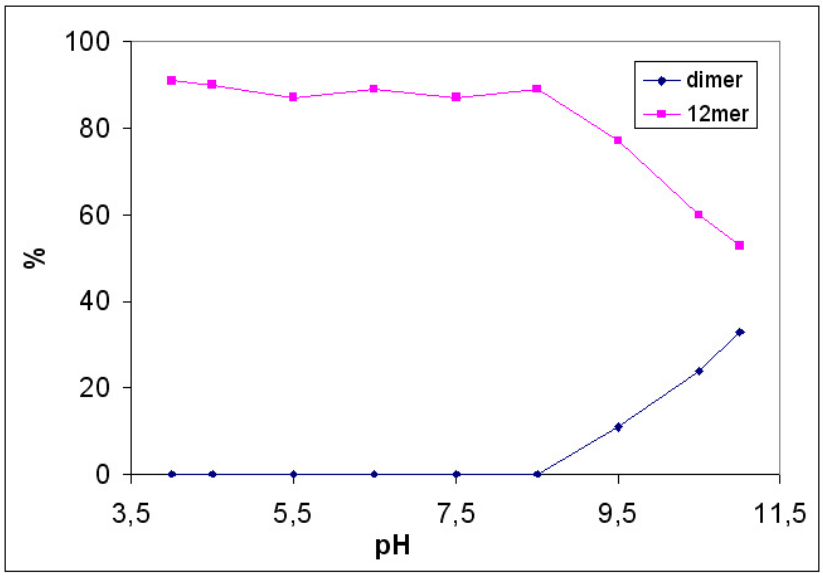


Figure 4.4: Oligomerization of PhTET3 with respect to pH. The dodecameric enzyme is dissociated by basic pH, but dissociation is less complete than in the case of EDTA. The percentages do not add up to 100% because low concentrations of other particles than dimers and dodecamers are detected in the solution.

buffer	Acetate		MES		PIPES	HEPES	TAPS	CHES	CAPS	
pH	4.0	4.5	5.5	6.5	7.5	7.5	8.5	9.5	10.5	11.0

Table 4.2: List of buffer systems used for testing PhTET3 oligomerization as a function of pH.

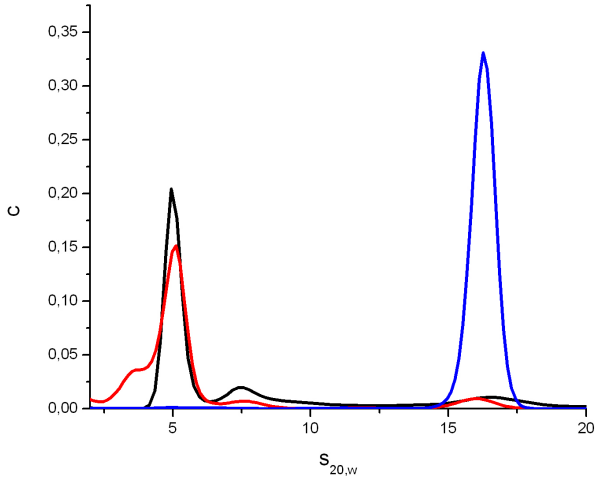


Figure 4.5: Thermostability of dimeric PhTET3. A dimeric sample was heated for two hours at 90 °C before EDTA was removed by dialysis (red line). Control1 (black line) shows a dimeric sample that was not heated before removal of EDTA and control2 (blue line) shows a dodecameric sample that was heated for two hours. The dimeric sample is less thermostable than the dodecameric one. Heating the dimer leads to monomer formation and reduces reoligomerization into dodecamers upon the removal of EDTA.

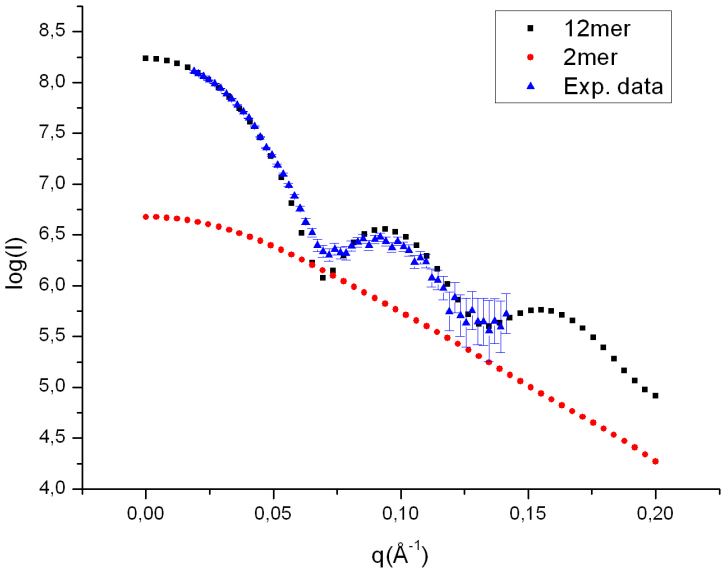


Figure 4.6: Theoretical scattering intensities for the PhTET3 dimer and dodecamer superposed with experimental data on a dodecameric sample. The theoretical curves are shown as they are calculated by CRYSON [134], whereas the experimental data is scaled by a constant so as to superpose it on the theoretical curve. The shape of the curves reflects the shape of the particles and the intensities at q=0 depend on the mass of the particles. The experimental data are in good agreement with the theoretical curve.

dodecamer. The shape of the curve for the dodecamer resembles the characteristic curve of a sphere. The profile of the curve is not surprising as the complex has a globular form. For the PhTET3 dodecamer the first minimum is located at $q = 0.07\text{Å}^{-1}$. The experimental data are in good agreement with the calculation, but have a smeared-out first minimum. This can be due to the sample not being perfectly monodisperse or to the finite resolution of the instrument. The dimer's calculated curve is flatter than that of the dodecamer, reflecting the smaller size of the particle. The first minimum of the curve of the dimer is outside the q-range shown in figure 4.6. The lower molecular weight of the dimeric particle leads to a lower intensity $I(0)$. The values for the radius of gyration R_g obtained from the calculations are 48Å for the dodecamer and 29Å for the dimer.

Thermal stability of the PhTET3 dodecamer. As pointed out in chapter 3, our SANS data show that the dodecameric complex is very thermostable. An experiment was carried out on the dodecameric complex in 20 mM Tris at pH 7.5 at 20 ℃ at a protein concentration of 2.3 mg/ mL. The results of the experiment are shown in figure 4.7. Furthermore we measured dodecameric PhTET3 in 20 mM Tris and 150 mM KCl at pH 7.5 at 90 ℃ after overnight incubation at the same temperature. The data are also shown in figure 4.7 and the parameters for each curve are summarized in table 4.3. From the similarity with the curve at 20 ℃ we conclude that the quaternary structure of PhTET3 has not changed.

SANS studies on PhTET3 dimer. After these preliminary experiments we measured a sample of PhTET3 in 20 mM Tris and 20 mM EDTA at 20 ℃ at a protein concentration of 2.3 mg/ mL. As can be seen in figure 4.7, the shape of the curve deviates from that of the dodecamer. A Guinier analysis resulted in a radius of gyration of 46 ± 0.5Å. This could indicate that the solution is a mixture of dodecameric and dimeric particles as the R_g corresponds neither to the dodecamer nor to the dimer. A little indentation at the position for the first minimum of the dodecameric particle can be seen. By comparing the extrapolated intensity of the direct beam of this sample with that of sample 1 and with the assumption that dimer and dodecamer are the only species present, the dimer content was evaluated to be 39%. The calculation can be found in A-1. To extract a low resolution structure from SAS data a monodisperse sample is required (see 2.3.2). Hence a fourth experiment was carried out in 20 mM Tris, 150 mM KCl and 20mM EDTA at 70 ℃. The data in figure 4.7 show that the curve for sample 4 is flatter than that of sample 3 indicating higher dimer content. However the $I(0)$ value indicates that the solution does not contain pure dimer. A similar calculation as for the third sample revealed a dimer content of 91%, indicating that the sample was still not sufficiently pure to get precise structural information. As a last attempt to obtain a pure sample of dimers at a concentration suitable for SANS experiments, PhTET3 was incubated in 20 mM EDTA overnight at 90 ℃. This incubation led to protein aggregation. Therefore we were not able to model the shape of the dimer from the SANS data.

4.4 Discussion

Our results clearly showed that the dodecameric assembly is very resistant against deoligomerization. Neither high salt concentration nor elevated temperature seem to perturb the quaternary structure. This is not surprising as

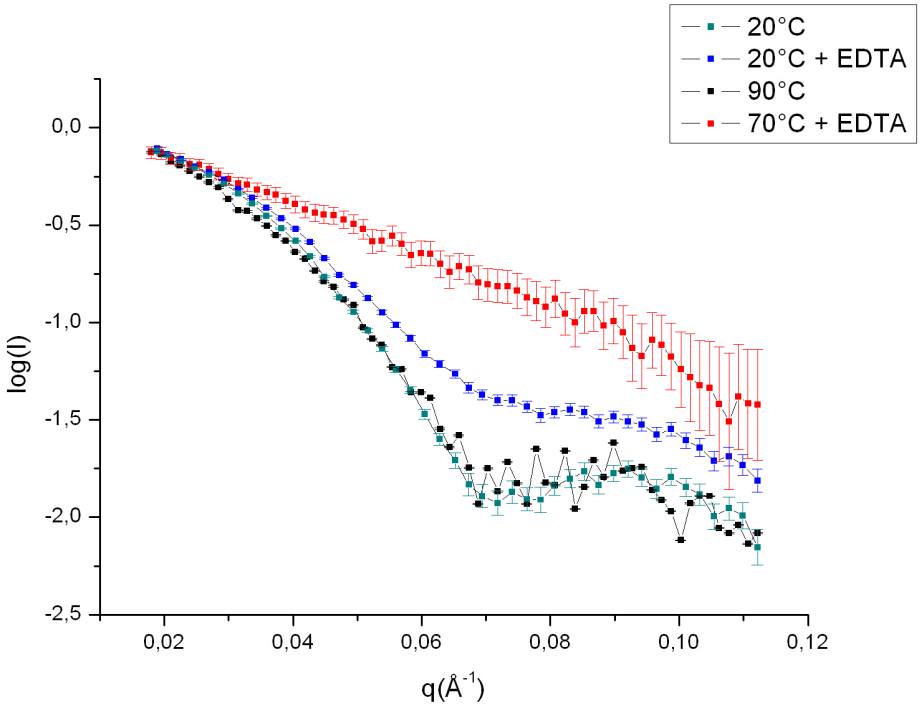


Figure 4.7: Experimental SANS data on PhTET3 normalized to $I(0)$. The samples at 20 ℃ and 90 ℃ both have the characteristic shape of the curve of the dodecamer with a minimum at $q = 0.07\text{Å}^{-1}$. When adding EDTA the minimum is less distinct at 20 ℃ and absent at 70 ℃. Characteristic parameters of the curves are shown in table 4.3.

	20 ℃	20 ℃ + EDTA	90 ℃	70 ℃ + EDTA
R_g (Å)	48.8 ± 0.8	46.5 ± 0.5	52 ± 1	38.5 ± 0.9
$I(0)$	1.01	0.684	0.715	0.137
Scaling factor	1.00	1.45	1.40	6.40
% dimer	0	39	0	90
Concentration(mg/ml)	2.3	2.3	3.1	3.1

Table 4.3: Summary of the parameters of SANS experiments on PhTET3 dimerization shown in figure 4.7. R_g and I_0 are the radius of gyration and the intensity at $q = 0$, calculated by using the Guinier approximation. The scaling factor corresponds to the constant with which the experimental data was multiplied in order to superpose the curves.

the protein is exposed to temperatures above 90 ℃ *in vivo*. However divalent cations appear to play an important role in the stabilization of the PhTET3 dodecamer. Similar observations have been reported for other multimeric enzymes [135, 136, 137, 138]. In the crystallographic structure of the complex the only divalent cations present are the two catalytic zinc ions in the active site. Therefore our results seem to suggest that the removal of the catalytic metal leads to deoligomerization of the dodecamer. Harmon et al. [135] reported the case of an enzyme from propionic acid bacteria which is stabilized against dissociation by Co^{2+} ions. In this case the cobalt responsible for the stabilization of the quaternary structure is not the same as the cobalt bound in the active sites, but stabilization is achieved by loosely bound cobalt that is in equilibrium with free cobalt ions in solution whose concentration must be maintained at 2 mM. Unfortunately Harmon et al. do not describe how this stabilization by cobalt works at an atomic level [135]. If a similar mechanism of stabilization by loosely bound ions would be responsible for oligomerization control in the case of PhTET3 this would explain why no other divalent cations are detected in the crystallographic structure. However as prolonged dialysis against buffer solutions fails to dissociate PhTET3 the stabilization by loosely bound ions seems improbable.
If PhTET3 deoligomerization is actually achieved by removing the catalytic ions, this mechanism seems improbable for oligomerization control *in vivo* as several days of dialysis against EDTA are required, indicating that the metal is tightly bound. Furthermore dimers obtained by treatment with EDTA are less thermostable than the dodecameric protein. As it is likely that dimeric PhTET3 exists *in vivo*, high thermostability of the dimer is essential.
In addition we observed that PhTET3 dissociates into dimers at basic pH. Our biochemical analysis (see chapter 3) showed that the enzyme is optimally active at pH 7.4. At pH 8.5 we did not detect any dimers, but the enzymatic activity only amounted to about 80% of the maximum activity. Hence the observed activity loss above pH 7.4 seems not to be caused by oligomer dissociation.
Our attempts to obtain structural information on the apo-dimer were not successful because the results obtained by AUC could not be reproduced at higher protein concentrations necessary for structural studies. Even at a protein concentration of 1mg/ml complete dimerization of the sample could not be achieved (data not shown). This suggests that the dodecameric form is stabilized by higher protein concentration.

Chapter 5

Activity and Stability of PhTET3 under high hydrostatic pressure

5.1 Introduction

A large portion of the earth's biosphere is a high pressure environment [139]. Around 79% of the marine biosphere lies at depths below 1000 m, i.e. at pressures of 100 bar or higher. Recent oceanographic campaigns revealed the existence of a large number of species in the deep-sea [75] **(more refs)**. It is well known that hydrostatic pressure shifts chemical equilibria towards the state with smaller volume. This affects all biochemical or physiological processes in an organism when it is exposed to high pressure. In organisms that are adapted to live at atmospheric pressure exposure to high pressure will lead to growth arrest and eventually to cell death [31]. On the other hand piezophilic organisms have been isolated that require elevated pressure for optimal growth [140]. The piezophilic bacteria and archaea that have been isolated so far are closely related to surface organisms which are pressure sensitive, showing that adaptation to high hydrostatic pressure does not require vast changes in the organism [139].
On a cellular level it has been observed that mesophilic microbes become filamentous when cultured under moderately high pressures that are below the pressure where growth is inhibited [139]. Interestingly when piezophiles are grown at ambient pressure they too become filamentous [139]. When exposed to high pressure *E. coli* bacteria express cold and heat shock proteins [141]. Pressure is the only known stressor to induce the expression of members of both groups of proteins simultaneously [139]. Indeed the effect of high pressure resembles that of high temperature as it destabilizes the quaternary and tertiary structure of proteins and it resembles that of low

temperatures since it has a similar effect on protein synthesis and membrane structure [139]. It has been observed that deep-sea organisms have a higher percentage of unsaturated fatty acids in their membranes [141]. This could be a way to maintain the viscosity of the membrane at a certain level or this could be related to the permeability of the membrane for certain ions [139]. Additionally transmembrane proteins are affected by high pressure. On the one hand changes in transmembrane-protein structure and activity could be due to a direct influence of high pressure on the protein, on the other hand membrane proteins could also be influenced by pressure induced changes in the surrounding lipids [141]. Moreover DNA synthesis has been shown to be one of the most pressure sensitive cellular processes during long-term pressure incubation [139]. Furthermore the dissociation of ribosomes could be a major factor in pressure induced cell death [141]. However Gross et al. have shown that functional ribosomes are stable in the physiological relevant pressure range *in vitro* [142]. Another known effect of moderately high pressure is the deoligomerization of protein complexes and inactivation of certain enzymes [73]. A hint that the growth arrest of microorganisms might be related to pressure induced loss of enzyme function is the fact that its kinetics resemble the inactivation kinetics of proteins [56]. Since the Ph-TET3 protease comes from a hyperthermophilic and piezotolerant species, it is expected to be able to function at high temperature and high pressure. As mentioned before in section 1.2.3, to fulfill their role in the organism, proteins do not only need to adopt a precise three-dimensional structure, but this structure also has to be sufficiently flexible to allow the conformational changes associated with the proteins physiological role. This is achieved by a delicate balance between stabilizing and repulsive interactions in the protein structure that control protein stability and flexibility. Studying the stability and the biological activity of proteins from organisms adapted to high temperature and high pressure thus may help to gain a better understanding of the interactions stabilizing proteins [143]. Besides TET proteases form big dodecameric complexes with large cavities inside (see section 1.3.2). This makes them good models for studying the stabilization mechanisms of high molecular weight complexes.

Furthermore enzymes that are active under high hydrostatic pressure have potential biotechnological applications. In 2006 Simonato et al. [141] wrote: "From a biotechnological point of view we consider piezophiles to be a resource that is awaiting development." In food biotechnology moderate hydrostatic pressure may be used to influence the metabolism of bacteria to reduce the formation of certain metabolites [144]. Furthermore the lethal effect of pressure on many microorganisms is used for food sterilization [141, 145]. Studying the influence of pressure on biological molecules and more specifi-

cally on large molecular assemblies will help to further develop this field of biotechnology.

5.2 Materials and Methods

5.2.1 Production of recombinant PhTET3

The protein was obtained as described in chapter 3.

5.2.2 Small Angle X-ray Scattering

All SAXS experiments were carried out at the high brilliance beamline ID02 at the European Synchrotron Radiation Facility in Grenoble, France. For the high pressure measurements the cell described in [146] was used. The sample to detector distance was set to 2m. The x-ray energy used was 16.5keV which corresponds to a wavelength of 0.751Å. All samples were measured at a concentration of 1mg/ml. Exposure times were 5 times 0.3s for the sample and 10 times 0.1s for the buffer. During data reduction the dark images were subtracted from the data, which were then corrected for spatial distortion and divided by a flatfield image. Data reduction was automatically done during the data acquisition at ID02. The program Fit2d [147] was used for masking of the beamstop, the detector edges and any parasitic scattering from the diamond windows of the pressure cell. Averaging, subtraction of the mask and azimuthal integration were done using the Saxs program package by P. Boesecke [148]. In a last step scaled buffer intensities were subtracted from the sample intensities using the program PRIMUS [133]. Pair distribution functions and the corresponding radii of gyration were calculated using the program GNOM [149].

5.2.3 Activity tests/enzymology

To determine aminopeptidase activity H-lysine-pNA[1] was used and product release was monitored by following the absorbance at 405 nm. Activity measurements under pressure were carried out on a photospectrometer described in [150] at INSERM U710, Montpellier, France using the same reaction buffer as at atmospheric pressure (see chapter 3) except that HEPES was used because its pH shows a lower variation with pressure than TAPS. The temperature was reduced to 50 °C to slow down the reaction. As the absorbance could be followed directly, the reaction was not stopped with acetic acid. No

[1] p-nitroanilide

mineral oil was used either, because the cuvette was closed by a film in direct contact with the reaction mixture.

5.3 Results

5.3.1 Small angle x-ray scattering under extreme conditions

To elucidate whether the PhTET3 dodecamer can exist under high hydrostatic pressure, we carried out small angle scattering experiments under high pressure. In contrast to the experiments described in section 4.3.2, the experiments were done using x-rays instead of neutrons. As the diamond windows of the high pressure cell scatter the incident beam, x-rays are better suited for this kind of experiment because a higher flow can be achieved compared to neutrons. Parasitic scattering from the pressure cell was masked from the scattering data. The parasitic scattering led to bad data quality at very small angles in some data sets. Therefore the pair distribution function was always used to calculate the radius of gyration. All experiments were done at a protein concentration of 1mg/ml in HEPES buffer. A test run showed that PhTET3 did not suffer measurable radiation damage when exposed to the beam for 130 s in intervals of 1 s or less.
Theoretical calculations similar to those described in section 4.3.2 were done to determine the theoretical radii of gyration for the dodecamer and dimer. We used the program CRYSOL [151] adapted for x-ray scattering experiments. The theoretical radius of gyration was calculated to be (50.5±0.5)Å for the dodecamer and (29.9±0.1)Å for the dimer. As explained in section 4.3.2, due to the lack of structural information on the dimer the theoretical calculations were performed on the dimer as it is in the dodecamer.

The PhTET3 dodecamer is not altered by very high hydrostatic pressure at ambient temperature. In a first experiment dodecameric PhTET3 was measured at 25 ℃ and at pH 7 at different hydrostatic pressures. The radius of gyration as a function of hydrostatic pressure is shown in figure 5.1. The radius of gyration is in agreement with the theoretical value and does not change under high pressure of up to 3000 bar. This means that the complex is not dissociated under pressure. In figure 5.2 a superposition of the scattering curves recorded at atmospheric pressure and at 3000 bar is shown. The shape of the curve does not change with pressure, indicating that no important structural changes are induced by high pressure.

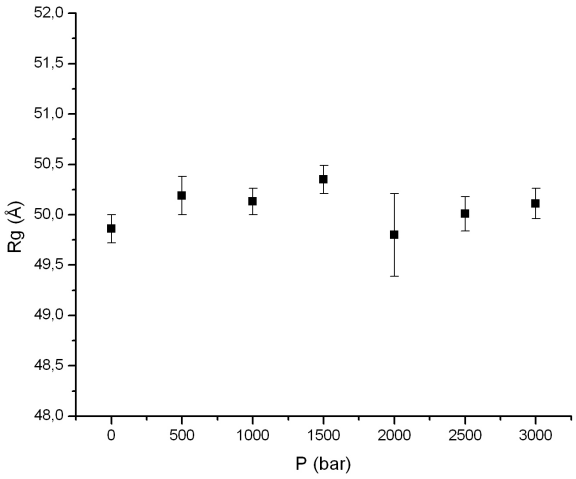


Figure 5.1: Radius of gyration of PhTET3 under high hydrostatic pressure at pH7 and at 25 °C

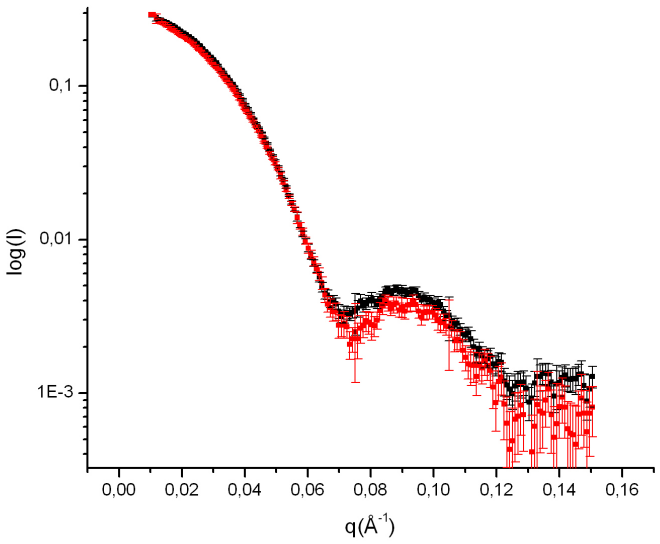
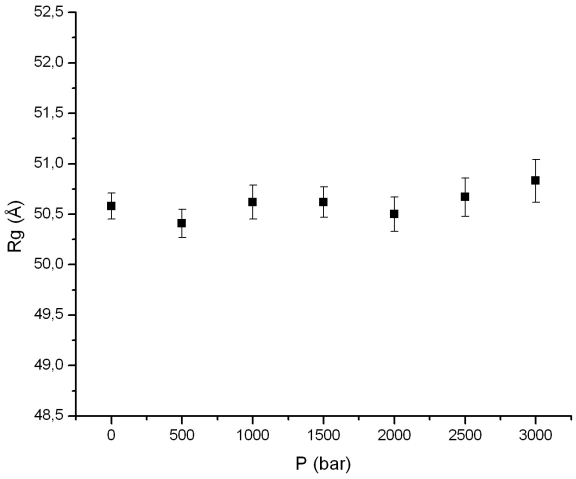


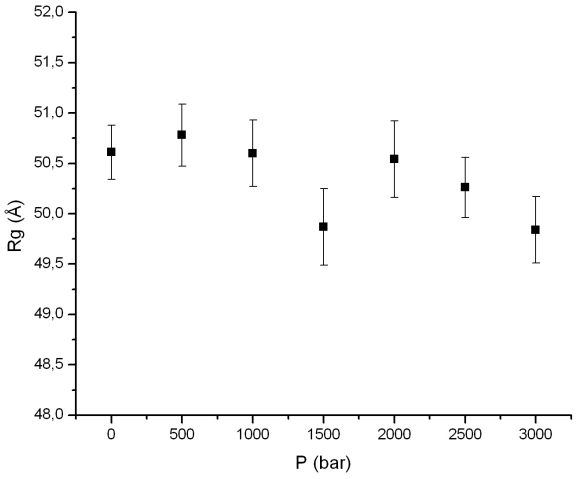
Figure 5.2: Superposition of scattering curves of PhTET3. The curves measured at pH 7 at 25 °C at atmospheric pressure (■) and at 3000 bar (■) are identical.

The pressure resistance of the PhTET3 dodecamer is not diminished by basic pH. In an attempt to further destabilize the assembly we raised the pH as we were able to show that PhTET3 is partly dissociated into dimers by basic pH (see section 4.3.1). Experiments were carried out at pH 9 and pH 10. Figure 5.3 shows the radius of gyration of PhTET3 under these two conditions. At ambient pressure the radius of gyration is in excellent agreement with the theoretical value. Furthermore it can be seen when superposing the curves measured at atmospheric pressure at basic pH with those measured at neutral pH that the shape of the curves does not change (data not shown). This means that during the SAXS experiments the oligomeric form of the protein is not altered by basic pH. As discussed in section 4.4 this is probably due to the PhTET3 dodecamer being stabilized by the higher protein concentration used in the SAXS experiments.
As the pressure is raised the radius of gyration does not change within the experimental error, showing that the PhTET3 dodecamer is not destabilized by hydrostatic pressure of up to 3000 bar even at basic pH. As at pH 7 the shape of the curve is not altered with high pressure (data not shown), indicating that no major conformational changes take place.

The combined effect of hydrostatic pressure and temperature does not deoligomerize the PhTET3 dodecamer. Furthermore we recorded scattering data on dodecameric PhTET3 at pH 7 and at 60 °C at different hydrostatic pressures. The results are shown in figure 5.4. At atmospheric pressure the radius of gyration is only slightly lower than the theoretical value. As the pressure is raised to 500 bar the radius of gyration is reduced by over 1Å. At first view it seems as if the radius of gyration would then increase with pressure from 500 bar to 3000 bar. However one should keep in mind that the resolution of small angle scattering data is limited and that data recorded in a high pressure cell is especially noisy. The difference between the smallest radius of gyration at 500 bar and the radius of gyration at 3000 bar is only 1.2Å which is probably not significant. In any case a difference of less than 2Å in the radius of gyration between the different pressures in this experiment and a deviation of less than 3Å from the theoretical value might indicate minor changes in the structure but this is not consistent with a deoligomerization of the particle. When superposing the data recorded at pH7 at 25 °C at atmospheric pressure with that recorded at pH 7 at 60 °C at 3000 bar (figure 5.5), it can be seen that the shape of the curves is identical within the experimental error. Above all the positions of the first minimum and of the first side maximum do not change, indicating that the shape of the particle is the same under the two conditions. This shows that the



(a) Radius of gyration of PhTET3 at pH9



(b) Radius of gyration of PhTET3 at pH10

Figure 5.3: Radius of gyration of PhTET3 as a function of high hydrostatic pressure at basic pH and at 25 °C.

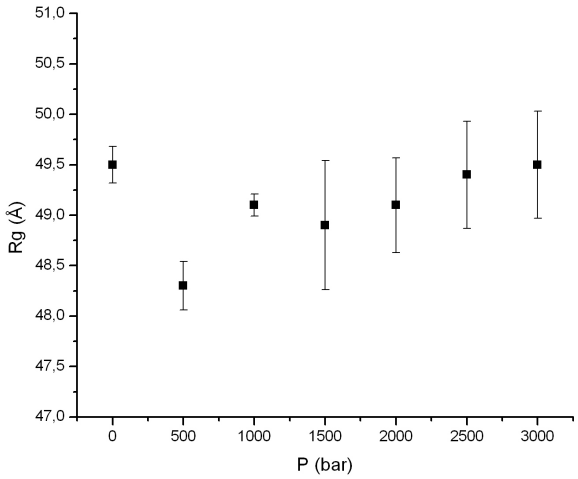


Figure 5.4: Radius of gyration of PhTET3 under high hydrostatic pressure at pH 7 and at 60 °C.

complex does not only stay in its dodecameric form, but that the exposure to high hydrostatic pressure and high temperature does not provoke important conformational changes.

The combined effect of hydrostatic pressure and temperature does not deoligomerize the PhTET3 dodecamer (2). In a last experiment a sample of dodecameric PhTET3 in a buffer that was adjusted to be at pH 9 at 90 °C was placed in the beam at ambient temperature. The pressure was then raised to 3000 bar and the sample was slowly heated up to 90 °C. While the sample was heated, scattering curves were recorded at four different temperatures. The radius of gyration at the different temperatures is shown in figure 5.6. Within the experimental error, the radius of gyration at 22 °C, 38 °C and 58 °C is in agreement with the theoretical value. At 86 °C the radius of gyration is (51.3±0.3)Å, which is slightly larger than the theoretical value. This could indicate a swelling of the dodecamer with temperature that could be related to a molten globule transition in the monomers, but as mentioned above this could also be due to the noisy data obtained with the high pressure cell. As there is no clear trend in the four data points available, the significance of the bigger radius of gyration at 86 °C is ambiguous. It is however unequivocal that the sample is dodecameric at 3000 bar and 90 °C. In figure 5.5 the scattering curve at 3000 bar and 86 °C is shown together with the data recorded at 25 °C at pH 7 and at atmospheric pressure. The curves

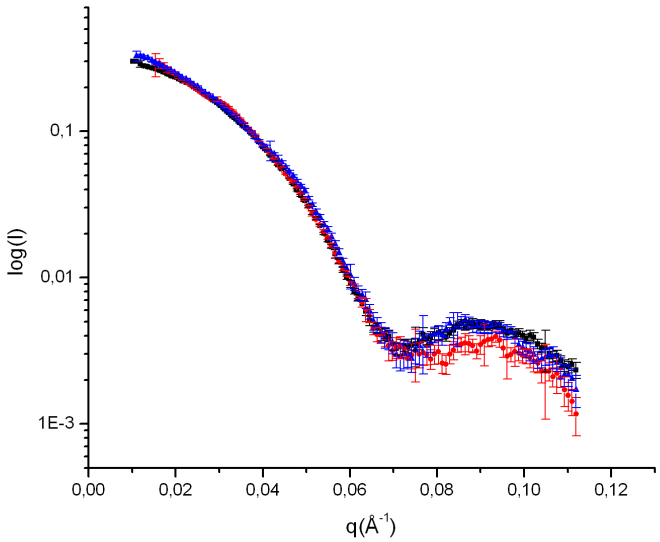


Figure 5.5: Superposition of scattering curves of PhTET3. The shape of the three curves, one at pH 7, 25 °C and atmospheric pressure (■), one at pH 7, 60 °C, 3000 bar (●) and one at pH 9, 90 °C, 3000 bar (▲), is identical.

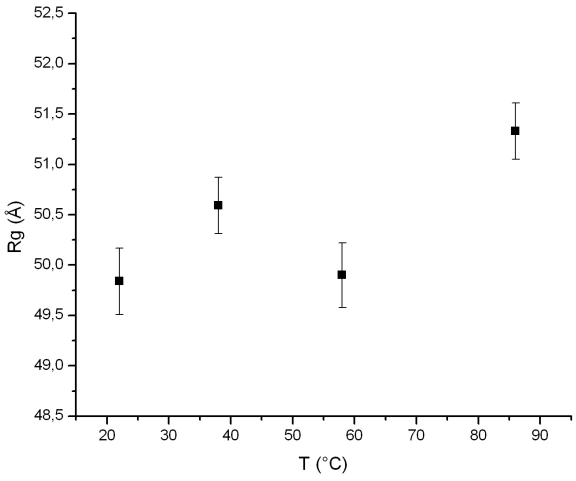


Figure 5.6: Radius of gyration of PhTET3 at 3000 bar at four different temperatures.

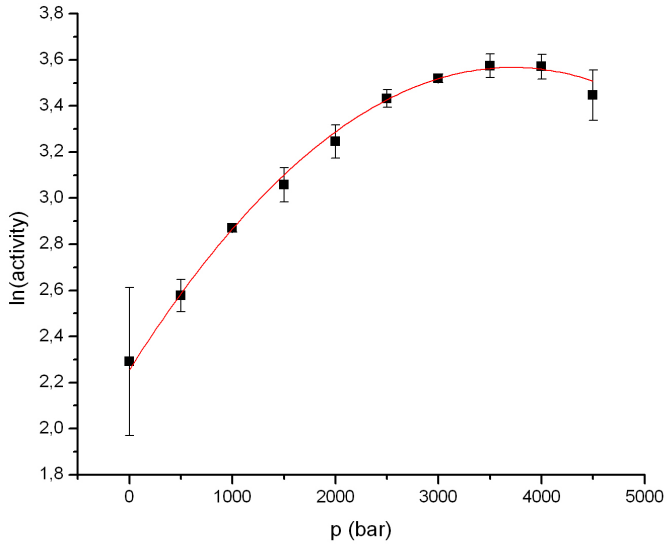


Figure 5.7: Activity of PhTET3 as a function of pressure. The experiment was carried out at a substrate concentration of 5 mM Lys-pNA and the activity is calculated as the number of micromoles of pNA that are liberated per milligram of PhTET3 per minute. The speed of the reaction increases as the pressure increases and is maximal between 3500 and 4000 bar.

superpose very well, showing that no major structural changes have occurred in the dodecamer between the two conditions. Furthermore incubation of the sample at 90 °C and 2000 bar for 1 hour did not lead to changes in the radius of gyration nor in the shape of the curve (data not shown).

5.3.2 Enzymatic activity under pressure

To elucidate whether the structural stability of PhTET3 under pressure is coupled with enzymatic activity and to assess the effect of hydrostatic pressure on enzyme catalysis, we studied the kinetic parameters of the hydrolysis of H-Lys-pNA by PhTET3 under pressure. To measure the speed of the reaction the amount of liberated pNA was measured in a spectrophotometer. As it takes a few minutes to load the sample into the pressure cell and to raise the pressure, the reaction had to be slowed down. Therefore all enzymatic tests were carried out at 50 °C. In figure 5.7 the influence of pressure on the enzymatic activity is shown. Duplicate measurements were made at each pressure. It can be seen that the activity of PhTET3 is enhanced by

	$\Delta V^{\neq}_{Vmax} = \Delta V^{\neq}_{kcat}$
Fit 1	
0-1500 bar	(-12 ± 6)ml/mol
1500-3250 bar	(4 ± 2)ml/mol
Fit 2	
0-750 bar	(-24 ± 11)ml/mol
750-3250 bar	(1 ± 3)ml/mol
	$-\Delta V_{Km} = \Delta V_{binding}$
0-3250bar	(-0.6 ± 0.2)ml/mol

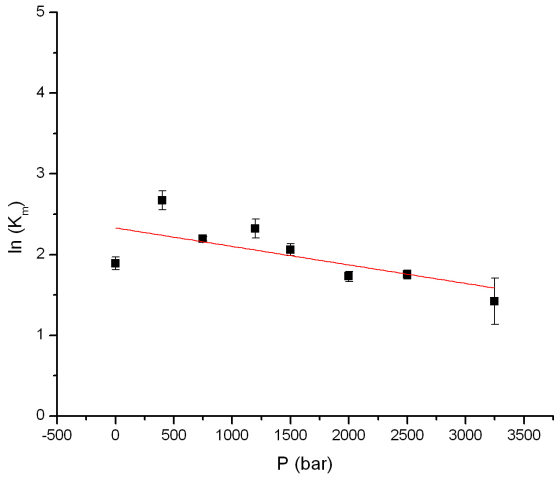
Table 5.1: Volume changes of steps in the hydrolysis of Lys-pNA by PhTET3, calculated from the slopes of the curves shown in figure 5.8

pressure and that the enzyme is active under very high hydrostatic pressure of up to 4500 bar.
Furthermore we determined the kinetic parameters of the reaction. To do this, duplicate essays were carried out at eight or more different substrate concentrations and the data were fitted with the Michaelis-Menten equation (equation 2.11) to extract the values of the Michaelis constant K_m and of the maximum reaction velocity V_{max} at each pressure. In figure 5.8 K_m and V_{max} are shown as a function of pressure. From the slope of the plots the corresponding volume changes ΔV were determined as described in section 1.2.3. The results are summarized in table 5.1. The data show that the volume change upon productive substrate binding is slightly negative, while the volume change of the catalytic step is negative at low pressure and takes a small positive value at higher pressure. Since reactions with negative volume changes are favored in a high pressure environment, this shows that substrate binding is always promoted by pressure whereas the catalytic step is enhanced at lower pressure and becomes unfavorable at higher pressure.

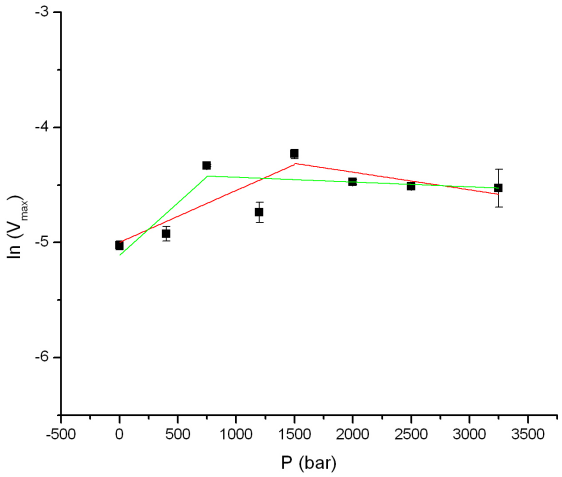
5.4 Discussion

5.4.1 Stability of PhTET3 quaternary structure under high pressure

The SAXS experiments clearly show that the structure and the oligomeric state of PhTET3 are maintained over a wide range of temperature and pressure. Since oligomeric proteins are known to be dissociated by high hydro-



(a) K_m as a function of pressure.



(b) V_{max} as a function of pressure.

Figure 5.8: K_m and V_{max} as a function of pressure. The data was fitted by linear regression. In the case of V_{max} the data are equivocal with respect to where the break in the fit should be. The two possible fits are shown.

static pressure [31] it is quite surprising to find that large complexes such as TET proteases can resist elevated pressure this well. Examples of studies on the quaternary structure of proteins are reported in section 1.2.3. Numerous oligomeric proteins studied under pressure dissociate at pressures lower than 3000 bar. In the case of GroEL from *E. coli* that has a similar number of subunits (14) as TET, the onset of dissociation was observed at a pressure of 1300 bar. However a pressure of 1300 bar corresponds to a depth of 13000 m in water, which is about the depth of the Mariana trench (11000 m), the deepest part of the ocean on earth. This is a very remarkable stability for a protein from a surface organism. The pressure stability of PhTET3 extends even beyond the pressures organisms could encounter at the deepest depth of the ocean. It is also interesting to see that the pressure resistance of PhTET3 extends over a wide range of temperatures. Therefore the question arises why PhTET3 displays this extraordinary stability. One possibility is that the pressure resistance of PhTET3 is linked to the architecture of TET proteases. In this case it should be displayed by all TET proteins. It is also possible that the pressure stability of PhTET3 is a side-effect of its hyperthermophilic character. A link between pressure and temperature resistance has been suggested before (see section 1.2.3). Finally the pressure stability of PhTET3 could be due to the adaptation of the source organism, *P. horikoshii*, to a high pressure habitat. It has been suggested that the continued exposure to high pressure *in vivo* could require protein stability beyond the physiologically relevant pressure range [72].

5.4.2 Enzymatic Activity of PhTET3 under high pressure

The hydrolysis of H-Lys-pNA by PhTET3 is enhanced by hydrostatic pressure in the physiologically relevant range and beyond. From the activity assays we were able to determine the volume changes associated with different steps in the catalytic mechanism. The negative volume change associated with substrate binding is maintained throughout the whole assayed pressure range. In contrast there is a break in the behavior of the volume change associated with the catalytic step which is negative at lower pressure and becomes positive as pressure is increased. As explained in section 1.2.3 this can have several reasons. Non-linear behavior can be due to deoligomerization or denaturation of the enzyme. In the case of PhTET3 this can be excluded because the SAXS data clearly show that the enzyme retains its oligomeric state in the relevant pressure range and that no major conformational changes take place. Another possible explanation for non-linear pressure dependence

of catalytic rate constants can be a change in the compressibility of the enzyme. The pressure dependence of PhTET3 at a non-saturating substrate concentration (5mM Lys-pNA, see figure 5.7) can be described very well by a quadratic fit. In non-saturating conditions the observed volume change $\Delta V^{\neq}_{obs}$ can be described as follows:

$$\Delta V^{\neq}_{obs} = \Delta V^{\neq}_{kcat} - \frac{\Delta V_{pb}}{1 + [S]/K_m}$$

A quadratic pressure dependence is typical for reactions that are influenced by compressibility changes in the protein. In this case $\Delta V^{\neq}_{obs}$ can be expressed as shown in equation 1.4. However since the observed pressure dependence of $\Delta V^{\neq}_{kcat}$ can best be approximated by two linear curves (see figure 5.8), we propose that the most probable explanation of the non-linear behavior is a change in the rate-limiting step. There are no data on the catalytic mechanism of TET proteases available, but it is thought to be identical to that of AAP [103]. In AAP it is assumed that the rate limiting step is the breaking of the C-N bond [115]. Furthermore it was shown that in AAP the rate-limiting step does not change with temperature from 25-85 ℃. It seems possible that product release becomes rate-limiting in the case of PhTET3 with higher pressure. We have proposed that the free amino acids generated by PhTET3 are expelled through the small opening in the apex of the tetrahedron (see chapter 3). As pressure is known to reduce the volume and the flexibility of proteins [82, 55], the mechanism of product release of PhTET3 could be disturbed by pressure. It is also possible that the release of free pNA is inhibited by pressure. As the pNA molecules are bulkier than lysines the compression of the protein could strongly affect its mobility inside the cavities. In addition a phenylalanine residue (F241) is found in the catalytic chamber about 10 Å from the active site. Since aromatic stacking interactions are known to be enhanced by pressure (see section 1.2.3), the release of pNA could be the new rate-limiting step.
The volume change associated with substrate binding in PhTET3 is relatively small. For other proteases such as thermolysin, volume changes of about -1 to -27ml/mol for substrate binding have been reported [152]. On the other hand the volume change associated with the catalytic step in PhTET3 in the physiologically relevant range (up to 1100 bar) is of the same order of magnitude as that observed for thermolysin (about +30ml/mol [152]). However the volume change observed for PhTET3 is negative, leading to an activation of the reaction under pressure. In contrast the volume change associated with the rate limiting step in catalysis at higher pressure is small compared to the values observed for thermolysin. Thus, even though the volume change is negative, catalysis should not be strongly inhibited by pressure.

5.4.3 Conclusions

Since PhTET3 is enzymatically active and oligomeric under hydrostatic pressure in the physiologically relevant range and beyond, the pressure tolerance of *P. horikoshii* seems not to be limited by inhibition of PhTET3. However when interpreting the activity data under pressure one should keep in mind that the volume changes of the reaction depend on the interactions between enzyme, substrate and the solvent molecules. This is nicely illustrated by the example of the 20S proteasome from *M. jannaschii*. It has been shown that the reaction rate of the hydrolysis of CBz-AAL-pNA increased 3.4-fold when the pressure was raised to 500 bar [143]. However the rate of hydrolysis of a longer peptide, insulin B_{ox}, has been shown to decrease with increasing pressure [153]. The authors proposed that the observed rate increase under pressure was due to a reduction of the overall volume of the transition state [153]. In contrast a longer substrate forms more interactions with the interior of the proteasome that are all influenced by pressure and thus the correct positioning of a longer substrate will be influenced more strongly by pressure [153]. Therefore testing the activity of PhTET3 under pressure with a different substrate could lead to a different pressure behavior.
From the data obtained on one TET protease it is not possible to conclude whether PhTET3 is specifically adapted to high pressure. In order to be able to answer this question studies on the structural stability and enzymatic activity of TET proteases from other organisms are needed. Studying the structure of other TET proteases and PhTET3 mutants under pressure would also allow gaining insight into the structural mechanisms stabilizing quaternary structure under high pressure.

Chapter 6

Conclusion and Perspectives

From a biophysical point of view PhTET3 has been extensively characterized. We were able to show that PhTET1, 2 and 3 are highly specialized aminopeptidases that are complementary in their substrate specificity. Working together these aminopeptidases can degrade almost any peptide into free amino acids. By comparing the structures of the three proteins we identified the structural features responsible for the differences in specificity. In addition the structural comparison yielded information on the mode of action of TET proteases. In order to validate our model we are trying to co-crystallize TET proteins with substrate analogs. Up to now these attempts were not successful.

It appears that the metal-ions in the active site are not only essential for catalytic activity but also for the oligomerization of the particle. In this respect it would be interesting to measure the metal content of the particle by flame emission spectroscopy. This would allow determining with certainty whether there are other metals bound to the particle that could be responsible for the stabilization of the structure. Furthermore this technique would permit to resolve whether the zinc ions present in the active site of recombinant PhTET3 are readily replaced by cobalt ions. In this context we also plan to measure the metal affinity of PhTET3 by titration experiments. The mechanism of cobalt activation of PhTET3, that was also observed for PhTET2 and PhTET1, is yet unknown.

Concerning the mechanism controlling TET oligomerization *in vivo*, further *in vivo* studies are required as it seems likely that it is mediated by other cellular components such as biological macromolecules (other proteins, sugars or nucleic acids) or by small organic molecules. Without more knowledge from *in vivo* studies, it seems difficult to further the understanding of PhTET3 oligomerization control by *in vitro* experiments.

The stability under high pressure and high temperature displayed by Ph-

TET3 is astonishing for an oligomeric assembly this large. From the crystallographic structure it was not evident why the complex is highly stable under extreme conditions of temperature and pressure. Apart from potential biotechnological applications this offers the possibility to gain new insights into the structural determinants for the stability of oligomeric protein complexes. This can be achieved through a comparison with homologs from surface organisms and from mesophiles. In order to carry out these studies the purification and characterization of suitable TET proteases has been initiated in our laboratory. By carefully selecting the proteins to study we hope to be able to identify the structural mechanisms responsible for the thermo- and baro-stability of PhTET3 and to discriminate between the effects of thermophilic and piezophilic adaptation. In order to assess whether the stability of PhTET3 is linked to its hyperthermophilic character, we aim to study a TET protease from a hyperthermophilic surface organism that is evolutionary distant from marine organisms. Only this evolutionary distance can ensure that the organism has not colonized the surface habitat only recently and thus ensure that a potential pressure stability of its TET protease is not a relic of a past in the deep-sea. In addition we plan to purify a mesophilic and a psychrophilic TET protease from surface organisms. When a suitable genome becomes available, we also hope to obtain a TET protease from a piezo-psychrophile.

By carrying out x-ray crystallography experiments under pressure, we aim to determine the influence of hydrostatic pressure on the structure of PhTET3 more precisely. We are especially interested in changes that could explain the effect of high pressure on the enzymatic activity of PhTET3. To this end it would also be useful to determine the high pressure structure of PhTET3 with an inhibitor bound to the active site. Besides, by closely monitoring structural changes on the subunit interfaces of different TET proteases that display distinct pressure stabilities, it should be possible to identify those residues responsible for the stability of the quaternary structure under pressure. Additionally mutation of appropriate residues in PhTET3 could be used to confirm findings from the comparison of homologs. Identifying a TET protease that is dissociated into dimers by high pressure would also allow to gain structural information on the TET dimer from small angle scattering data.

Finally we also intend to study the dynamics under pressure of whole cells by elastic incoherent neutron scattering (EINS). It has been shown that the major contribution to the scattering of cells comes from proteins [154]. Therefore it is possible to determine the mean flexibility and resilience of the proteins in the cell. Such measurements have already been successfully used in the study of thermo-adaptation [154]. The advantage of these measurements on

whole cells is that they assess the average properties of all proteins from an organism. Since different proteins from the same organism can have different strategies of adaptation, the study of whole cells can give a more general picture than the study of individual proteins.

Chapter 7

Resumé de la thèse en français

7.1 Introduction

7.1.1 Les archaeas

En 1990 Woese et al. [2] ont proposé l'introduction d'un arbre phylogénétique comprenant trois domaines de la vie : les bactéries, les archaeas et les eucaryotes. La relation entre ces trois domaines de la vie a été déterminée par analyses de séquences des ANRs ribosomales. Le domaine des archaeas a été identifié par Woese et Fox en 1977 [3]. Comme les bactéries, les archaeas sont des procaryotes mais les deux groupes sont suffisamment différents pour être considérés comme deux domaines. De nouveaux développements techniques en métagénomique ont permis d'étudier les communautés de microorganismes non-cultivables. Il a été démontré récemment que les archaeas sont nombreuses dans la plupart des écosystèmes mondiaux. Cependant les organismes adaptés aux conditions les plus extrêmes font partie des archaeas.
Dans l'arbre phylogénétique de Woese et al. [2], une racine est placée entre les domaines des archaeas et des bactéries. Cette racine représente le dernier ancêtre commun à tous les être vivants (LUCA, the last universal common ancestor). L'existence et la nature de LUCA sont fortement disputés. Des études géophysiques, phylogénétiques et de comparaison de génomes suggèrent que le LUCA aurait été adapté pour la vie à haute température et à haute pression [13, 12, 14, 17, 22, 26]. Afin d'avancer notre compréhension sur l'origine de la vie et sur la possibilité que ceci avait lieu dans de telles conditions il est important d'étudier les macromolécules issues des organismes hyperthermophiles et barophiles.
Dans le cadre de cette thèse nous avons étudié une protéine issue de l'archaea hyperthermophile *Pyrococcus horikoshii*. Elle a été isolée dans la fosse marine d'Okinawa dans le Pacifique à une profondeur de 1395m [28]. En plus d'être

hyperthermophile, *P. horikoshii* montre des caractéristiques barophiles ou barotolerantes en fonction des conditions de culture [29].

7.1.2 Adaptation aux conditions extrêmes

7.1.2.1 Les extremophiles

Sur terre il existe de nombreux endroits où règnent des conditions qui sont extrêmes d'un point de vue humain. Un exemple sont les océans dans lesquels on trouve de nombreux habitats de haute pression, de basse ou de haute température et de salinité élevée [30, 31]. Afin de vivre dans ces conditions déstabilisantes pour les macromolécules biologiques, les organismes ont développé des stratégies d'adaptation. Certains organismes sont capables d'empêcher le facteur déstabilisant d'entrer dans le cytoplasme et d'autres ont développé des méthodes pour les expulser. Mais aussi certains organismes se sont adaptés en transformant leurs macromolécules de façon à être stables dans les conditions extrêmes. Cependant les mécanismes stabilisant les protéines en conditions extrêmes ne sont pas bien connus [37] surtout pour les assemblages oligomériques. Néanmoins on sait que les protéines des extremophiles sont faites à partir des vingt acides aminés connus et que leur stabilité extraordinaire est due à un nombre d'interactions stabilisant plus élevé [37, 79]. Cette augmentation du nombre d'interactions stabilisantes est atteinte à travers des changements subtils dans la composition en acides aminés de la protéine.

L'adaptation à la haute température En comparant les structures tridimensionnelles de protéines homologues issues d'organismes adaptés à différentes températures, des stratégies d'adaptation à la température ont été identifiées. Les modifications observées incluent l'optimisation du cœur hydrophobe, des interactions entre sous-unités, l'augmentation du nombre de liaisons hydrogène et de ponts salins et la diminution de la taille des boucles mobiles [33, 39, 41, 42]. Toutefois il est connu que la fonctionnalité d'une enzyme ne dépend pas que de sa structure mais aussi de son comportement dynamique. Des experiences de diffusion de neutrons ont démontré que les protéines hyperthermophiles et mésophiles ont une flexibilité similaire à leurs températures d'activité maximale, mais que la protéine hyperthermophile est plus rigide [51].

L'adaptation à la haute pression Une grande partie de la biosphère est un environnement de haute pression. L'effet de la pression sur un système suit le principe de Le Chatelier, i.e. l'état avec le plus petit volume va être favorisé. Pour un système en équilibre entre deux états, caractérisé par une

constante κ (qui correspond à K pour le cas statique et à k pour le cas dynamique), celle-ci est liée à la différence de volume ΔV entre les deux états comme décrit par l'équation 7.1 [55].

$$\left(\frac{\partial(ln\kappa)}{\partial P}\right)_T = \frac{-\Delta V}{RT} \tag{7.1}$$

Dans cette équation R dénote la constante des gaz parfaits et T la température. La différence de volume est constituée du changement de volume de la protéine et d'une contribution du solvant dû à l'influence de la pression sur l'hydratation de la protéine. Par ailleurs des effets dus aux co-solvants peuvent se produire. En général c'est la contribution du solvant qui contribue le plus sur la différence de volume [54].
Tandis que la structure primaire d'une protéine est essentiellement insensible à la pression dans la gamme physiologique et au delà [54], les interactions faibles stabilisant les structures secondaires, tertiaires et quaternaires sont affectées par la pression [50, 54, 57]. Les interactions ioniques sont déstabilisées par la pression à cause de la réduction de la densité du solvant par électrostriction autour de charges exposées [31, 54, 57, 58]. Le même effet a été attribué à la solvatation des surfaces hydrophobes entrainant la déstabilisation de ces interactions par la pression [31, 54, 60]. Cependant il existe des rapports sur la stabilisation des interactions hydrophobes par la pression [79, 50]. Les liaisons hydrogène sont principalement insensibles à la pression [31, 54, 59] et l'empilement d'anneaux aromatiques est favorisé par la pression [54]. En général les protéines monomériques sont assez résistantes à la pression hydrostatique et ne sont dénaturées qu'au delà de pressions de 3000 à 4000 bar, mais la dénaturation à des pressions plus basses a été observée [13, 31, 68, 69]. Cependant les protéines oligomériques peuvent être dissociées par des pressions plus basses, qui sont dans la gamme physiologique de pression [13, 30].
L'activité des enzymes est affectée par la pression en fonction du changement de volume qui est associé à la réaction. Ce changement de volume ΔV peut être déterminé avec l'équation 7.1 en étudiant le comportement de K_m et k_{cat} sous pression. A partir de K_m on calcule le changement de volume lors de la liaison du substrat ΔV_{pb} et k_{cat} donne accès au volume d'activation $\Delta V^{\neq}$. Des réactions avec des changements de volume positifs sont inhibés par la pression et ceux avec un ΔV negatif sont favorisés. Dans la plupart des cas des comportements linéaires sont observés, mais des effets non-linéaires ont déjà été vus. Trois effets pouvant expliquer ce comportement ont été identifiés : un changement dans l'étape limitante de la réaction, un changement dans la compressibilité de la protéine ou un changement dans la structure de la protéine.

Des études sur l'adaptation des protéines à la pression ont été faites sur des enzymes de poissons. Dans ces études les effets de la pression sur l'activité d'enzymes issues de poissons de la mer profonde ont été comparés à ceux sur l'activité d'homologues de poissons vivants à la surface. Pour des lactate deshydrogenases (LDH) une baisse de l'activité enzymatique avec la pression a été détectée, mais cette baisse était moins forte pour les LDHs des poissons abyssaux [76]. Le changement de volume observé pendant la polymérisation de l'actine est plus petit dans la protéine issue des poissons vivants à des pressions plus hautes [71]. Ces résultats indiquent qu'il y a une adaptation à la pression au niveau des protéines. Toutefois les mécanismes structuraux responsables de cette adaptation ne sont pas connus [72]. Le fait que des organismes vivant dans des environnements marins de haute pression doivent aussi être adaptés à la haute ou la basse température présente une difficulté additionnelle [31]. Vu qu'un grand nombre de protéines thermostables montre une grande stabilité sous haute pression, il a été suggéré qu'il existe un lien entre l'adaptation à la pression et à la haute température [79, 46, 80].
Afin de mieux comprendre la baroadaptation et les mécanismes de stabilisation des protéines, des études supplémentaires semblent nécessaire.

7.1.3 La protéolyse

7.1.3.1 Importance et mécanisme général

Dernièrement il a été réalisé que la dégradation des protéines dans la cellule est un processus essentiel pour le fonctionnement des mécanismes moléculaires [83]. La protéolyse intervient dans des processus aussi divers que le cycle cellulaire, la régulation des facteurs de transcription, la présentation des antigènes et dans des voies métaboliques [84]. Par ailleurs la protéolyse est impliquée dans la dégradation des protéines mal repliées [83]. Des facteurs externes comme la haute température, des concentrations de sel élevées ou la présence de petites molécules réactives peuvent augmenter le nombre de protéines mal repliées [85]. Ainsi la protéolyse est d'autant plus importante dans des conditions extrêmes [85, 87].
Dans les trois domaines de la vie, la protéolyse dans le cytosol repose principalement sur des grands assemblages macromoléculaires autocompartimentés. Afin d'éviter la dégradation non-spécifique des protéines, les sites actifs de ces grandes protéases se trouvent dans des cavités à l'intérieur des molécules [85]. L'accès à ces cavités est limité aux protéines dépliées [84]. Le complexe du proteasome est le membre le plus connu de cette famille de protéases. Le proteasome a été identifié dans les eucaryotes, les archaeas et dans certaines bactéries [90]. Dans des bactéries qui ne possèdent pas de proteasome, des

autres complexes comme ClpP, HslV ou Lon ont été découverts [90]. Ces protéases qui utilisent de l'énergie pour la dégradation des protéines, produisent des polypeptides d'une longueur entre trois et 30 acides aminés [92]. Il a été proposé que ces polypeptides soient ensuite dégradés par des protéases auto-compartementalisées, qui ne nécessitent pas d'énergie [91]. Les protéases Tricorn, TOP[1], Gal6 / bleomycin hydrolase, DppA et TPPII ont été proposées comme faisant partie de ce groupe [94, 96, 97, 98, 99]. Il a été montré que dans des cellules dans lesquelles le proteasome est inhibé, l'activité de ces protéases, indépendante de l'énergie, était augmentée. Ceci pourrait indiquer que ces protéases permettent à la cellule de compenser la perte de l'activité du proteasome [99, 100]. Bien qu'un grand nombre de protéases ait été caractérisé, leur rôle dans la cellule et les mécanismes de régulation ne sont pas encore bien connus. Pour cela il est important d'étudier d'autres protéases *in vitro* et *in vivo* afin de mieux comprendre cette fonction essentielle.

7.1.3.2 Les protéases TET

Un nouveau groupe de protéases auto-compartementalisées a été découvert dans *Haloarcula marismortui* par Franzetti et al. [1]. Contrairement aux complexes de protéases connus ce nouveau groupe ne forme pas de complexes en forme de "tonneau", mais en forme de tétraèdre [103]. Ces protéases ont alors été nommées TET [1]. Les protéases TET font partie de la famille M42 selon la classification MEROPS [104]. Des membres de cette famille ont été identifiés dans une grande partie des génomes d'archaea séquencés à ce jour [105]. Parmi ces génomes il n'y a qu'un seul qui comporte la protéase Tricorn et une protéase M42 [105]. Ainsi il a été suggéré que Tricorn et TET pourraient être des homologues fonctionnels [1, 106].
Quatre structures de protéases de la famille M42 ont été déterminées par cristallographie aux rayons X, dont deux appartenant à *P. horikoshii*. Les deux protéases de *P. horikoshii* ont été identifiées comme des protéases TET [108, 103, 107]. Il a été démontré que les TET sont des aminopeptidases [1, 106]. L'architecture du site actif et la classification de MEROPS suggèrent que le mécanisme catalytique des TET ressemble à celui d'AAP [103]. Les structures de TET publiées montrent une particule homo-dodécamerique, auto-compartementalisée [1, 103, 108, 107]. Chaque monomère de TET est composé de deux domaines : un domaine catalytique et un domaine de dimérisation. L'élément de base pour former le dodécamère est vraisemblablement un dimère. Dans le complexe dodécamérique chaque arête est formée d'un dimère. Sur chaque face du tétraèdre trois dimères sont assemblés en un

[1]thimet oligopeptidase

cercle avec un pore d'entrée au milieu qui donne accès à la cavité centrale du complexe avec les sites actifs localisés près des apex du tétraèdre.

7.1.4 Objectifs de la thèse

Pourquoi il y a-t-il plusieurs protéases TET dans certains organismes ? Bien que les protéases TET ont été caractérisées *in vitro*, leur rôle *in vivo* n'est pas connu. Par ailleurs dans certains organismes comme *H. marismortui* il n' y a qu'une protéase TET, tandis que dans d'autres organismes comme *P. horikoshii* il y en a plusieurs. Les structures de deux protéases TET de *P. horikoshii* (PhTET1 et PhTET2) ont été résolues et la biochimie de PhTET2 a été étudiée [108, 103, 107, 106]. En examinant la structure et l'activité enzymatique d'une troisième protéase de la famille M42 de *P. horikoshii*, PhTET3, *in vitro* nous espérons trouver la raison de l'existence de plusieurs protéases TET dans certains organismes et de déterminer un rôle potentiel pour les protéases TET *in vivo*.

Comment l'oligomerisation de PhTET3 est-elle contrôlée *in vivo* ? Il a été montré que l'oligomerisation et l'auto-compartementalisation jouent un rôle important dans le fonctionnement et la régulation des protéases dans le cytoplasme. En utilisant PhTET3 comme modèle nous étudions l'assemblage oligomérique des protéases TET. Nous voulons identifier l'unité de base du tétraèdre et caractériser sa structure. En plus nous déterminons la stabilité de l'oligomère par rapport à des facteurs chimiques comme les métaux, le sel et le pH.

Comment se comporte l'oligomère PhTET3 en conditions pseudo-physiologiques ? *P. horikoshii* est une archaea hyperthermophile capable de pousser à des pressions élevées. La haute température et la haute pression sont connues pour déstabiliser les complexes oligomériques. Conséquemment il semble possible que PhTET3 soit peu stable *in vivo*. Nous étudions donc la structure et l'activité enzymatique de PhTET3 en conditions pseudo-physiologiques, i.e. haute température et haute pression.

7.2 Résultats

7.2.1 Caractérisation de PhTET3

Nous avons caractérisé un complexe de peptidase de 468-KDa de l'archaea hyperthermophile *Pyrococcus horikoshii*. La structure de la protéine a été

déterminée à une résolution de 1.9Å. Le monomère contient un domaine catalytique typique de la famille des peptidases M42 et un domaine de dimérisation qui permet la formation des dimères qui s'assemblent en tétraèdres auto-compartimentés avec 12 sous-unités, ressemblants à ceux décrits pour les peptidases TET. L'analyse biochimique montre que l'enzyme est activée par le cobalt et coupe des peptides par un mécanisme non-processif. Par conséquent, cette protéine représente un troisième complexe de peptidase TET décrit dans *P. horikoshii*, appelé de ce fait PhTET3. C'est une lysyl-aminopeptidase avec une préférence forte pour les résidus basiques, qui sont mal clivés par PhTET1 et PhTET2. L'analyse structurale de PhTET3 et sa comparaison avec PhTET1 et PhTET2 révèlent les dispositifs communs expliquant le mode d'action des machines moléculaires TET, aussi bien que des différences qui peuvent être associées à la discrimination forte de substrat. La question de la stabilité des TETs sous des températures extrêmes a été abordée. PhTET3 montre une activité maximale à 95°C et des expériences de diffusion de neutrons aux petits angles à 90°C démontrent l'absence de changements de structure quaternaire après des temps prolongés d'incubation. En conclusion, les PhTETs forment un ensemble complémentaire de machines de destruction de peptides qui peut jouer un rôle important dans le métabolisme de *P. horikoshii*.

7.2.2 Oligomerisation de PhTET3

Dans les archaeas halophiles il n'existe qu'une protéase TET. Des études sur *Halobacterium salinarium* ont montré que TET est présente sous une forme dimérique et sous une forme dodécamérique *in vivo* (résultats non publiés). Quand la PhTET3 recombinante est purifiée on ne trouve qu'une forme dodécamérique. En vue des résultats des études *in vivo* il semble probable qu'il existe aussi différents états oligomériques de PhTET3 *in vivo*. Toutefois, comme le rôle de PhTET3 dans la cellule n'est pas connu, on ignore la raison pour l'existence du dimère. Afin de mieux comprendre l'oligomerisation de PhTET3 nous avons étudié la stabilité du dodécamère *in vitro*.
En utilisant la centrifugation analytique et la diffusion des neutrons aux petits angles nous avons montré que le dodécamère de PhTET3 est très stable. Ni des concentrations de sel élevées, ni la haute température désoligomérisent le complexe. Cependant les cations divalents semblent jouer un rôle important dans la stabilisation du dodécamère, vu que nous avons pu observer que le complexe se dissocie en dimères en présence d'EDTA. Dans la structure cristallographique les seuls cations divalents détectés sont ceux du site actif. Ainsi nos résultats suggèrent que l'élimination des métaux catalytiques entraine la désoligomérisation de la particule. Dès que l'EDTA est enlevé, une

partie des dimères se réoligomerise en dodecamères. La réoligomérisation est plus importante si on ajoute des métaux, mais dans tous les cas testés la majorité de la protéine reste sous forme de dimère. Les dimères obtenus ainsi sont moins thermostables que la protéine oligomérique. Comme la protéine provient d'un organise hyperthermophile il semble improbable que de tels dimères puissent exister *in vivo*. Un effet de désoligomérisation, moins fort que pour l'EDTA, a aussi été détecté en pH basique. Au delà de pH 8 une partie de la protéine est sous forme dimérique.
En utilisant l'EDTA pour désoligomériser PhTET3 nous avons essayé d'obtenir une structure à basse résolution du dimère par diffusion aux petits angles. Cependant nous n'avons pas réussi à obtenir une solution de dimère pure à une concentration en protéine suffisante pour les études biophysiques.

7.2.3 Activité et stabilité de PhTET3 sous pression hydrostatique

Afin de remplir leurs fonctions dans l'organisme, les protéines ont besoin d'adopter une conformation tridimensionnelle précise. En même temps la structure doit être suffisamment flexible pour permettre les changements conformationnels necessaires au rôle physiologique. Ceci est atteint par une balance délicate entre des interactions stabilisantes et déstabilisantes. Etudier la stabilité et l'activité de protéines adaptées à des hautes températures et à des hautes pressions peut aider à mieux comprendre les mécanismes de stabilisation des protéines. Etant donné que la protéase PhTET3 est issue d'un organisme hyperthermophile et barotolérant, on s'attend à ce qu'elle soit stable et active à haute température et à haute pression. De plus cette protéine forme de grands complexes macromoléculaires, ce qui fait d'elle une bonne protéine modèle pour ce genre d'études.
Nous avons utilisé la diffusion des rayons X à petits angles et des tests d'activité pour étudier le comportement de PhTET3 sous pression. Nos résultats montrent que la protéine reste oligomérique sous des conditions de hautes pressions (3000 bar) et de hautes températures (90 °C), même à pH basique (pH9). Nous avons également montré que PhTET3 est activée par la pression et qu'elle le reste jusqu'à des pressions d'au moins 4500 bar. De plus nous avons déterminé le K_m et le k_{cat} en fonction de la pression. Ceci nous a permis de calculer le volume d'activation et le changement du volume associé à la liaison du substrat. ΔV_{pb} est de -0.6ml/mol sur toute la gamme de pression testée (de la pression atmosphérique jusqu'à 3250 bar). La liaison du substrat est alors favorisée par la pression. Le volume d'activation est d'abord négatif et devient légèrement positif à haute pression. Nous avons

attribué ce comportement à un changement d'étape limitante de la réaction. Nos résultats montrent que PhTET3 est active et stable dans les conditions pseudo-physiologiques pour *P. horikoshii*.

7.3 Conclusions

D'un point de vue biophysique PhTET3 a été intensivement caractérisé. Nous avons montré que PhTET1, 2 et 3 sont des aminopeptidases hautement spécialisées qui sont complémentaires dans leurs spécificités de substrat. En collaboration ces amniopeptidases peuvent dégrader presque n'importe quel peptide en acides aminés libres. En plus en comparant les structures des trois protéines nous avons identifié les dispositifs structuraux responsables des différences dans la spécificité.
En ce qui concerne le mécanisme contrôlant l'oligomérisation de TET *in vivo*, d'autres études *in vivo* sont requises car il semble probable que ce contrôle est effectué par l'intermédiaire d'autres composants cellulaires tels que les macromolécules biologiques (d'autres protéines, sucres ou acides nucléiques) ou des petites molécules organiques. Sans plus de connaissance provenant d'études *in vivo*, il semble difficile de faire avancer la compréhension du mécanisme contrôlant l'oligomérisation de PhTET3 par des expériences *in vitro*.
La stabilité sous pression et à hautes températures montrée par PhTET3 est étonnante pour un assemblage oligomérique de cette taille. En plus des applications biotechnologiques potentielles, ceci offre la possibilité d'étudier les déterminants structuraux pour la stabilité des complexes oligomériques par une comparaison avec des homologues d'organismes de surface et mésophiles. Au moyen de la cristallographie aux rayons X sous pression, nous visons à déterminer l'influence de la pression hydrostatique sur la structure de PhTET3 avec plus de précision. Nous sommes particulièrement intéressés par des effets qui pourraient expliquer l'effet de la pression sur l'activité enzymatique de PhTET3.
Enfin nous avons l'intention d'étudier la dynamique sous pression des cellules entières par la diffusion incohérente élastique de neutrons (EINS). L'avantage de ces mesures sur des cellules entières est qu'elles évaluent les propriétés dynamiques moyennes de toutes les protéines d'un organisme. Puisque des protéines différentes du même organisme peuvent avoir différentes stratégies d'adaptation, l'étude des cellules entières peut donner une vue plus générale que l'étude de protéines individuelles.

Bibliography

[1] B. Franzetti, G. Schoehn, J. F. Hernandez, M. Jaquinod, R. W. Ruigrok, and G. Zaccai. Tetrahedral aminopeptidase: a novel large protease complex from archaea. *EMBO J*, 21(9):2132–2138, 2002.

[2] C. R. Woese, O. Kandler, and M. L. Wheelis. Towards a natural system of organisms: Proposal for the three domains archaea, bacteria and eucarya. *Proc Natl Acad Sci U S A.*, 87:4576–4579, 1990.

[3] Carl R. Woese and George E. Fox. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc Natl Acad Sci U S A.*, 74(11):5088–5090, 1977.

[4] J.R. Brown and W.F. Doolittle. Archaea and the prokaryote-to-eukaryote transition. *Microbiol Mol Biol Rev.*, 61(4):456–502, 1997.

[5] C. Woese. The universal ancestor. *Proc Natl Acad Sci U S A.*, 95(12):6854–6859, 1998.

[6] B. Chaban, S.Y.M. Ng, and K.F. Jarrell. Archaeal habitats - from the extreme to the ordinary. *Can J Microbiol*, 52:73–116, 2006.

[7] E.F. DeLong. Everything in moderation: Archaea as "non-extremophiles". *Curr Opin Genet Dev*, 8:649–54, 1998.

[8] S.G. Tringe, C. von Mering, A. Kobayashi, A.A. Salamov, K. Chen, H.W. Chang, M. Podar, J.M. Short, E.J. Mathur, J.C. Detter, P. Bork, P. Hugenholtz, and E.M. Rubin. Comparative metagenomics of microbial communities. *Science*, 308(5721):554–557, 2005.

[9] E.V. Pikuta, R.B. Hoover, and J. Tang. Microbial extremophiles at the limits of life. *Crit Rev Microbiol*, 33(3):183–209, 2007.

[10] E.J. Javaux. Extreme life on earth - past, present and possibly beyond. *Res Microbiol*, 157:37–48, 2006.

[11] S. Gribaldo and C. Brochier-Armanet. The origin and evolution of archaea: a state of the art. *Phil. Trans. R. Soc. B*, 361:1007–1022, 2006.

[12] E. G. Nisbet and N. H. Sleep. The habitat and nature of early life. *Nature*, 409(6823):1083–1091, 2001.

[13] I. Daniel, P. Oger, and R. Winter. Origins of life and biochemistry under high-pressure conditions. *Chemical Society Reviews*, 35(10):858–875, 2006.

[14] K.O. Stetter. Hyperthermophiles in the history of life. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1474):1837–1843, 2006.

[15] R.J.C. Hennet, N.G. Holm, and M.H. Engel. Abiotic synthesis of amino acids under hydrothermal conditions and the origin of life: a perpetual phenomenon? *Naturwissenschaften*, 79:361–365, 1992.

[16] U. Müller. Re-creating an RNA world. *Cellular and Molecular Life Sciences (CMLS)*, 63(11):1278–1293, 2006.

[17] D. W. Schwartzman and C. H. Lineweaver. The hyperthermophilic origin of life revisited. *Biochem. Soc. Trans.*, 32(Pt 2):168–171, 2004.

[18] V. Moulton, P.P. Gardner, R.F. Pointon, L.K. Creamer, G.B. Jameson, and D. Penny. Rna folding argues against a hot-start origin of life. *J Mol Evol*, 51:416–421, 2000.

[19] M. Tehei, B. Franzetti, M.-C. Maurel, J. Vergne, C. Hountondji, and G. Zaccai. The search for traces of life: the protective effect of salt on biological macromolecules. *Extremophiles*, 6(5):427–430, 2002.

[20] P.L. Knauth. Salinity history of the earth's early ocean. *Nature*, 395(6702):554, 1998.

[21] P. Forterre, S. Gribaldo, and C. Brochier. Luca : à la recherche du plus proche ancêtre commun universel. *M/S : médecine sciences*, 21(10):860–865, 2005.

[22] M. Di Giulio. The universal ancestor was a thermophile or a hyperthermophile. *Gene*, 281(1-2):11–17, 2001.

[23] N. R. Pace. A Molecular View of Microbial Diversity and the Biosphere. *Science*, 276(5313):734–740, 1997.

[24] W. F. Doolittle and J. R. Brown. Tempo, mode, the progenote, and the universal root. *Proceedings of the National Academy of Sciences of the United States of America*, 91(15):6721–6728, 1994.

[25] M. Di Giulio. The late stage of genetic code structuring took place at a high temperature. *Gene*, 261(1):189–195, 2000.

[26] M. Di Giulio. The ocean abysses witnessed the origin of the genetic code. *Gene*, 346:7–12, 2005.

[27] M. Di Giulio. The universal ancestor and the ancestor of bacteria were hyperthermophiles. *Journal of Molecular Evolution*, 57(6):721–730, 2003.

[28] J.M. González, Y. Masuchi, F.T. Robb, J.W. Ammerman, D.L. Maeder, M. Yanagibayashi, J. Tamaoka, and C. Kato. Pyrococcus horikoshii sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. *Extremophiles*, 2(2):123–130, 1998.

[29] K. Horikoshi. Barophiles: deep-sea microorganisms adapted to an extreme environment. *Curr. Opin. Microbiol.*, 1(3):291–295, 1998.

[30] R. Jaenicke. Enzymes under extremes of physical conditions. *Annu Rev Biophys Bioeng*, 10:1–67, 1981.

[31] M. Gross and R. Jaenicke. Proteins under pressure. the influence of high hydrostatic pressure on structure, function and assembly of proteins and protein complexes. *Eur J Biochem*, 221(2):617–630, 1994.

[32] L.J. Rothschild and R.L. Mancinelli. Life in extreme environments. *Nature*, 409(6823):1092–1101, 2001.

[33] R. Jaenicke. What ultrastable globular proteins teach us about protein stabilization. *Biochemistry (Mosc)*, 63(3):312–321, 1998.

[34] R. Ladenstein and G. Antranikian. Proteins from hyperthermophiles: stability and enzymatic catalysis close to the boiling point of water. *Adv Biochem Eng Biotechnol*, 61:37–85, 1998.

[35] D. Madern, C. Ebel, and G. Zaccai. Halophilic adaptation of enzymes. *Extremophiles*, 4(2):91–98, 2000.

[36] L. Costenaro, G. Zaccai, and C. Ebel. Link between protein solvent and weak protein protein interactions gives insight into halophilic adaptation. *Biochemistry*, 41:13245–13252, 2002.

[37] R. Jaenicke. Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? *PNAS*, 97(7):2962–2964, 2000.

[38] N. Sinha and S.J. Smith-Gill. Protein structure to function via dynamics. *Protein Pept Lett*, 9(5):367–77, 2002.

[39] R. Jaenicke and G. Böhm. The stability of proteins in extreme environments. *Curr Opin Struct Biol.*, 8(6):738–748, 1998.

[40] R. Jaenicke. Protein stability and molecular adaptation to extreme conditions. *Eur J Biochem*, 202(3):715–728, 1991.

[41] C. Vieille and G.J. Zeikus. Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev.*, 65(1):1–43, 2001.

[42] G.N. Somero. Adaptation of enzymes to temperature: searching for basic "strategies". *Comp. Biochem. Physiol. B*, 139(3):321–333, 2004.

[43] M. Sadeghi, H. Naderi-Manesh, M. Zarrabi, and B. Ranjbar. Effective factors in thermostability of thermophilic proteins. *Biophys Chem.*, 119(3):256–270, 2006.

[44] P.A. Fields. Protein function at thermal extremes: balancing stability and flexibility. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.*, 129(2-3):417–431, 2001.

[45] L.D. Unsworth, J. van der Oost, and S. Koutsopoulos. Hyperthermophilic enzymes - stability, activity and implementation strategies for high temperature applications. *FEBS journal*, 274:4044–4056, 2007.

[46] E. Mombelli, E. Shehi, P. Fusi, and P. Tortora. Exploring hyperthermophilic proteins under pressure: theoretical aspects and experimental findings. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1595(1-2):392–396, 2002.

[47] A. Irimia, F.M.D. Vellieux, D. Madern, G. Zaccai, A. Karshikoff, G. Tibbelin, R. Ladenstein, T. Lien, and N.-K. Birkeland. The 2.9åresolution crystal structure of malate dehydrogenase from *Archaeoglobus fulgidus*: Mechanisms of oligomerisation and thermal stabilition. *J. Mol. Biol.*, 335:343–356, 2004.

[48] G.N. Somero. Protein adaptations to temperature and pressure: complementary roles of adaptive changes in amino acid sequence and internal milieu. *Comp. Biochem. Physiol. B*, 136(4):577–591, 2003.

[49] M. Tehei and G. Zaccai. Adaptation to high temperatures through macromolecular dynamics by neutron scattering. *FEBS Journal*, 274:4034–4043, 2007.

[50] D. J. Hei and D. S. Clark. Pressure stabilization of proteins from extreme thermophiles. *Appl Environ Microbiol*, 60(3):932–939, 1994.

[51] M. Tehei, D. Madern, B. Franzetti, and G. Zaccai. Neutron scattering reveals the dynamic basis of protein adaptation to extreme temperature. *J Biol Chem*, 280(49):40974–40979, 2005.

[52] Jannasch H.W. and C.D. Taylor. Deep-sea microbiology. *Annu Rev Microbiol*, 38:487–514, 1984.

[53] G. N. Somero. Adaptions to high hydrostatic pressure. *Annu. Rev. Physiol.*, 54:557–577, 1992.

[54] V. V. Mozhaev, K. Heremans, J. Frank, P. Masson, and C. Balny. High pressure effects on protein structure and function. *Proteins*, 24(1):81–91, 1996.

[55] P. Masson and C. Balny. Linear and non-linear pressure dependence of enzyme catalytic parameters. *Biochim Biophys Acta*, 1724(3):440–450, 2005.

[56] K. Heremans and L. Smeller. Protein structure and dynamics at high pressure. *Biochim Biophys Acta*, 1386(2):353–370, 1998.

[57] K. Heremans. High pressure effects on proteins and other biomolecules. *Ann Rev Biophys Bioeng*, 11:1–21, 1982.

[58] B. B. Boonyaratanakornkit, C. B. Park, and D. S. Clark. Pressure effects on intra- and intermolecular interactions within proteins. *Biochim Biophys Acta*, 1595(1-2):235–249, 2002.

[59] P. C. Michels, D. Hei, and D. S. Clark. Pressure effects on enzyme activity and stability at high temperatures. *Adv Protein Chem*, 48:341–376, 1996.

[60] A. N. McCarthy and J. R. Grigera. Pressure denaturation of apomyoglobin: A molecular dynamics simulation study. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, 1764(3):506–515, 2006.

[61] G. Hummer, S. Garde, A. E. Garca, M. E. Paulaitis, and L. R. Pratt. The pressure dependence of hydrophobic interactions is consistent with the observed pressure denaturation of proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 95(4):1552–1555, 1998.

[62] K. Ado and Y. Taniguchi. Pressure effects on the structure and function of human thioredoxin. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, 1774(7):813–821, 2007.

[63] L. Mitra, N. Smolin, R. Ravindra, C. Royer, and R. Winter. Pressure perturbation calorimetric studies of the solvation properties and the thermal unfolding of proteins in solution—experiments and theoretical interpretation. *Physical Chemistry Chemical Physics*, 8:1249–1265, 2006.

[64] L.-N. Lin, J.F. Brandts, J.M. Brandts, and V. Plotnikov. Determination of the volumetric properties of proteins and other solutes using pressure perturbation calorimetry. *Analytical Biochemistry*, 302:144–160, 2002.

[65] F. Meersman, L. Smeller, and K. Heremans. Protein stability and dynamics in the pressure-temperature plane. *Biochim Biophys Acta*, 1764(3):346–354, 2006.

[66] Boris Gorovits, C. S. Raman, and Paul M. Horowitz. High Hydrostatic Pressure Induces the Dissociation of cpn60 Tetradecamers and Reveals a Plasticity of the Monomers. *J. Biol. Chem.*, 270(5):2061–2066, 1995.

[67] L. Erijman, G. H. Lorimer, and G. Weber. Reversible dissociation and conformational stability of dimeric ribulose bisphosphate carboxylase. *Biochemistry*, 32(19):5187–5195, 1993.

[68] J. L. Silva, D. Foguel, A. T. Da Poian, and P. E. Prevelige. The use of hydrostatic pressure as a tool to study viruses and other macromolecular assemblages. *Current Opinion in Structural Biology*, 6(2):166–175, 1996.

[69] P. Cioni and G. B. Strambini. Pressure effects on the structure of oligomeric proteins prior to subunit dissociation. *J. Mol. Biol.*, 263:789–799, 1996.

[70] M. Di Giulio. A comparison of proteins from *Pyrococcus furiosus* and *Pyrococcus abyssi*: barophily in the physicochemical properties of animo acids and in the genetic code. *Gene*, 346:1–6, 2005.

[71] T. Morita. Structure-based analysis of high pressure adaption of α-actin. *The Journal of Biological Chemistry*, 278(30):28060–28066, 2003.

[72] A. A. Brindley, R. W. Pickersgill, J. C. Partridge, D. J. Dunstan, D. M. Hunt, and M. J. Warren. Enzyme sequence and its relationship to hyperbaric stability of artificial and natural fish lactate dehydrogenases. *PLoS ONE*, 3(4):e2042, 2008.

[73] M. Gross, G. Auerbach, and R. Jaenicke. The catalytic activities of monomeric enzymes show complex pressure dependence. *FEBS Lett*, 321(2-3):256–260, 1993.

[74] E. Occhipinti, N. Bec, B. Gambirasio, G. Baietta, P. L. Martelli, R. Casadio, C. Balny, R. Lange, and P. Tortora. Pressure and temperature as tools for investigating the role of individual non-covalent interactions in enzymatic reactions Sulfolobus solfataricus carboxypeptidase as a model enzyme. *Biochim Biophys Acta*, 1764(3):563–572, 2006.

[75] J.S. Lipp, Y. Morono, F. Inagaki, and K.-U. Hinrichs. Significant contribution of archaea to extant biomass in marine subsurface sediments. *Nature*, 2008 (in press).

[76] Y. Nishiguchi, T. Miwa, and F. Abe. Pressure-adaptive differences in lactate dehydrogenases of three hagfishes: *Eptatretus burgeri, Paramyxine atami* and *Eptatretus okinoseanus*. *Extremophiles*, 12(3):477–480, 2008.

[77] M. M. Sun, R. Caillot, G. Mak, F. T. Robb, and D. S. Clark. Mechanism of pressure-induced thermostabilization of proteins: studies of glutamate dehydrogenases from the hyperthermophile *Thermococcus litoralis*. *Protein Sci*, 10(9):1750–1757, 2001.

[78] K. Muller, T. Seifert, and R. Jaenicke. High pressure dissociation of lactate dehydrogenase from *Bacillus stearothermophilus* and reconstitution of the enzyme after denaturation in 6 M guanidine hydrochloride. *Eur Biophys J*, 11(2):87–94, 1984.

[79] F. T. Robb and D. S. Clark. Adaptation of proteins from hyperthermophiles to high pressure and high temperature. *J Mol Microbiol Biotechnol*, 1(1):101–105, 1999.

[80] L. Smeller. Pressure-temperature phase diagrams of biomolecules. *Biochim Biophys Acta*, 1595(1-2):11–29, 2002.

[81] H. Ikeuchi, S. Kunugi, and K. Oda. Activity and stability of a neutral protease from *Vibrio* sp. (vimelysin) in a pressure-temperature gradient. *European Journal of Biochemistry*, 267(4):979–983, 2000.

[82] C.A. Royer. Insights into the role of hydration in protein structure and stability obtained through hydrostatic pressure studies. *Braz J Med Biol Res*, 38(8):1167–1173, 2005.

[83] A Varshavsky. Regulated protein degradation. *Trends in Biochemical Sciences*, 30(6):283–286, 2005.

[84] A. Ciechanover. Intracellular protein degradation: from a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Exp. Biol. Med.*, 231(7):1197–1211, 2006.

[85] A. Goldberg. Protein degradation and protection against misfolded proteins. *Nature*, 426:895–899, 2003.

[86] J.A. Maupin-Furlow, H.L. Wilson, S.J. Kaczowka, and M.S. Ou. Proteasomes in the archaea: from structure to function. *Front. bioscience*, 5:D837–865, 2000.

[87] R. Jaenicke, H. Schurig, N. Beaucamp, and R. Ostendorp. Structure and stability of hyperstable proteins: glycolytic enzymes from hyperthermophilic bacterium thermotoga maritima. *Adv Protein Chem*, 48:181–269, 1996.

[88] L. Vandeputten-Rutten and P. Gros. Novel proteases: common themes and surprising features. *Curr Opinion in Struct Biol*, 12:704–708, 2002.

[89] C.M. Pickart and R.E. Cohen. Proteasomes and their kin: proteases in the machine age. *Nature Reviews, Molecular Cell Biology*, 5:177–187, 2004.

[90] R. De Mot, I. Nagy, J. Walz, and W. Baumeister. Proteasomes and other selfcompartmentalizing proteases in prokaryotes. *Trends In Microbiology*, 7(2):88–92, 1999.

[91] C. N. Larsen and D. Finley. Protein translocation channels in the proteasome and other proteases. *Cell*, 91:431–434, 1997.

[92] A.F. Kisselev, T.N. Akopian, and A.L. Goldberg. Range of sizes of peptide products generated during degradation of different proteins by archaeal proteasomes. *J Biol Chem.*, 273(4):1982–1989, 1998.

[93] A.F. Kisselev, T.N. Akopian, K.M. Woo, and A.L. Goldberg. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J Biol Chem.*, 274(6):3363–3371, 1999.

[94] T. Tamura, N. Tamura, Z. Cejka, R. Hegerl, F. Lottspeich, and W. Baumeister. Tricorn protease - the core of a modular proteolytic system. *Science*, 274:1385–1389, 1996.

[95] N. Tamura, F. Lottspeich, W. Baumann, and T. Tamura. The role of tricorn protease and its aminopeptidase-interacting factors in cellular protein degradation. *Cell*, 95:637–648, 1998.

[96] T. Saric, C.I. Graef, and A.L. Goldberg. Pathway for degradation of peptides generated by proteasomes. *J Biol Chem*, 279(45):46723–46732, 2004.

[97] L. Joshua-Tor, H. E. Xu, S. A. Johnston, and D. C. Rees. Crystal Structure of a Conserved Protease That Binds DNA: The Bleomycin Hydrolase, Gal6. *Science*, 269:945–950, 1995.

[98] H. Remaut, C. Bompard-Gilles, C. Goffin, J.-M. Frère, and J. Van Beeumen. Structure of the *Bacillus subtilis* D-animopeptidase DppA reveals a novel self-compartmentalizing protease. *Natrue Structural Biology*, 8(8):674–678, 2001.

[99] E. Geier, G. Pfeiffer, M. Wilm, M. Lucchiari-Hartz, W. Baumeister, K. Eichmann, and G. Niedermann. A giant protease with potential to substitute for some functions of the proteasome. *Science*, 283:978–981, 1999.

[100] R. Glas, M. Bogyo, J.S. McMaster, M. Gaczynska, and H. L. Ploegh. A proteolytic system that compensates for loss of proteasome function. *Nature*, 392:618–621, 1998.

[101] E.W. Wang, B.M. Kessler, A. Borodovsky, B.F. Cravatt, M. Bogyo, H.L. Ploegh, and R. Glas. Integration of the ubiquitin-proteasome

pathway with a cytosolic oligopeptidase activity. *Proc Natl Acad Sci U S A*, 97(18):9990–9995, 2000.

[102] D. E. Ward, K. R. Shockley, L. S. Chang, R. D. Levy, J. K. Michel, S. B. Conners, and R. M. Kelly. Proteolysis in hyperthermophilic microorganisms. *Archaea*, 1:63–74, 2002.

[103] S. Russo and U. Baumann. Crystal structure of a dodecameric tetrahedral-shaped aminopeptidase. *J Biol Chem*, 279(49):51275–51281, 2004.

[104] Neil D. Rawlings and Alan J. Barrett. Merops: the peptidase database. *Nucleic Acids Res*, 27(1):325–331, 1999.

[105] http://merops.sanger.ac.uk.

[106] M. A. Dura, V. Receveur-Brechot, J. P. Andrieu, C. Ebel, G. Schoehn, A. Roussel, and B. Franzetti. Characterization of a TET-like aminopeptidase complex from the hyperthermophilic archaeon Pyrococcus horikoshii. *Biochemistry*, 44(9):3477–86, 2005.

[107] G. Schoehn, F. M. Vellieux, A. M. Dura, V. Receveur-Brechot, C. M. Fabry, R. W. Ruigrok, C. Ebel, A. Roussel, and B. Franzetti. An archaeal peptidase assembles into two different quaternary structures: A tetrahedron and a giant octahedron. *J Biol Chem*, 281(47):36327–36337, 2006.

[108] L. Borissenko and M. Groll. Crystal structure of TET protease reveals complementary protein degradation pathways in prokaryotes. *J Mol Biol*, 346(5):1207–1219, 2005.

[109] G. Chen, T. Edwards, V.M. D'souza, and R.C. Holz. Mechanistic Studies on the Aminopeptidase from *Aeromonas proteolytica*: A Two-Metal Ion Mechanism for Peptide Hydrolysis. *Biochemistry*, 36(14):4278–4286, 1997.

[110] C. Stamper, B. Bennett, T. Edwards, R.C. Holz, D. Ringe, and G. Petsko. Inhibition of the Aminopeptidase from *Aeromonas proteolytica* by L-Leucinephosphonic Acid. Spectroscopic and Crystallographic Characterization of the Transition State of Peptide Hydrolysis. *Biochemistry*, 40(24):7035–7046, 2001.

[111] C.C. De Paola, B. Bennett, R.C. Holz, D. Ringe, and G.A. Petsko. 1-Butaneboronic Acid Binding to *Aeromonas proteolytica* Aminopeptidase: A Case of Arrested Development. *Biochemistry*, 38(28):9048–9053, 1999.

[112] K. P. Bzymek and R. C. Holz. The Catalytic Role of Glutamate 151 in the Leucine Aminopeptidase from *Aeromonas proteolytica*. *J. Biol. Chem.*, 279(30):31018–31025, 2004.

[113] B. Bennett and R.C. Holz. Spectroscopically Distinct Cobalt(II) Sites in Heterodimetallic Forms of the Aminopeptidase from *Aeromonas proteolytica*: Characterization of Substrate Binding. *Biochemistry*, 36(32):9837–9846, 1997.

[114] B. Chevrier, C. Schalk, H. D'Orchymont, J.-M. Rondeau, D. Moras, and C. Tarnus. Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure*, 2(4):283–291, 1994.

[115] R.C. Holz. The aminopeptidase from *Aeromonas proteolytica*: structure and mechanism of co-catalytic metal centers involved in peptide hydrolysis. *Coord Chem Rev*, 232(1-2):5–26, 2002.

[116] G. Schurer, H. Lanig, and T. Clark. *Aeromonas proteolytica* aminopeptidase: An investigation of the mode of action using a quantum mechanical/molecular mechanical approach. *Biochemistry*, 43(18):5414–5427, 2004.

[117] S. Richard, F. Bonneté, O. Dym, and Zaccai G. The mpd-nacl-h2o system for the crystallisation of halophilic proteins. In F.T. Robb, A.R. Place, K.R. Sowers, H.J. Schreier, S. DasSarma, and E.M. Fleischmann, editors, *Archaea: a Laboratory Manual*, pages 149–154. Cold Spring Harbor Laboratory Press, New York, 1995.

[118] M.V. Petoukhov and D.I. Svergun. Analysis of x-ray and neutron scattering from biomacromolecular solutions. *Curr Opin Struct Biol*, 17(5):562–571, 2007.

[119] B. Jacrot and G. Zaccai. Determination of molecular weight by neutron scattering. *Biopolymers*, 20(11):2413–2426, 1981.

[120] A. Guinier. La diffraction des rayons x aux très faibles angles: applications à l'étude des phénomènes ultra-microscopiques. *Ann Phys (Paris)*, 12:161–236, 1939.

[121] http://www.embl hamburg.de/ExternalInfo/Research/Sax/software.html. Data analysis software atsas 2.2.

[122] L. Michaelis and M. Menten. Die Kinetik der Invertinwirkung. *Biochemische Zeitung*, 49:333–369, 1913.

[123] K. Schachmann. Ultracentrifugation, diffusion, and viscometry. *Methods in Enzymology*, 4:32–103, 1957.

[124] O. Lamm. Die differentialgleichung der ultrazentrifugierung. *Arkiv für matematik, astronomi och fysik*, 21B(2):1–4, 1929.

[125] P. Schuck. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J.*, 78(3):1606–1619, 2000.

[126] http://www.brenda enzymes.info/.

[127] N. J. Marianayagam, M Sunde, and J. M. Matthews. The power of two: protein dimerization in biology. *Trends in Biochemical Sciences*, 29(11):618–625, 2004.

[128] D. S. Goodsell and A. J. Olson. Structural symmetry and protein function. *Annu. Rev. Biomol.Struct.*, 29:105–153, 2000.

[129] I. M. A. Nooren and J. M. Thornton. Diversity of protein-protein interactions. *EMBO Journal*, 22(14):3486–3492, 2003.

[130] S. G. Rudyak, M. Brenowitz, and T. E. Shrader. Mg2+-Linked Oligomerization Modulates the Catalytic Activiy of the Lon (La) Protease from *Mycobacterium smegmatis*. 40(31):9317–9323, 2001.

[131] http://www.jphilo.mailway.com.

[132] R.E. Gosh, S.U. Egelhaaf, and A.R. Rennie. A computing guide for small-angle scattering experiments. *Institut Laue Langevin Internal Publications*, ILL06GH05T, 2006.

[133] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. J. Koch, and D. I. Svergun. *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography*, 36(5):1277–1282, 2003.

[134] D. Svergun, S. Richard, M.H.J. Koch, Z. Sayers, S. Kuprin, and G. Zaccai. Protein hydration in solution: Experimental observation by x-ray and neutron scattering. *Proc Natl Acad Sci USA*, 95(5):2267–2272, 1998.

[135] F. R. Harmon, N. H. Goss, and H. G. Wood. Stabilization of the Quaternary Structure of Transcarboxylase by Cobalt(II) Ions. *Biochemistry*, 21:2847–2852, 1982.

[136] S. P. Lee, J. Xiao, J. R. Knutson, M. S. Lewis, and M. K. Han. Zn^{2+} promotes the self-association of human immunodeficiency virus type-1 integrase *in vitro*. *Biochemistry*, 36:173–180, 1997.

[137] S. M. Green, A. Ginsburg, M. S. Lewis, and P Hensley. Roles of metal ions in the maintenance of the tertiary and quaternary structure of arginase from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, 266(32):21474–21481, 1991.

[138] M. Montenarh and R. Henning. Self-Assembly of Simian Virus 40 Large T Antigen Oligomers by Divalent Cations. *Journal of Virology*, 45(2):531–538, 1983.

[139] D. H. Bartlett. Pressure effects on *in vivo* microbial processes. *Biochim Biophys Acta*, 1595(1-2):367–381, 2002.

[140] R. Margesin and Y. Nogi. Psychropiezophilic microorganisms. *Cellular and Molecular Biology*, 50(4):429–436, 2004.

[141] F. Simonato, S. Campanaro, F.M. Lauro, A. Vezzi, M. D'Angelo, N. Vitulo, G. Valle, and D.H. Bartlett. Piezophilic adaptation: a genomic point of view. *J Biotechnol*, 126(1):11–25, 2006.

[142] M. Gross, K. Lehle, R. Jaenicke, and K. H. Nierhaus. Pressure-induced dissociation of ribosomes and elongation cycle intermediates. Stabilizing conditions and identification of the most sensitive functional state. *Eur J Biochem*, 218(2):463–468, 1993.

[143] P. C. Michels and D. S. Clark. Pressure-enhanced activity and stability of a hyperthermophilic protease from a deep-sea methanogen. *Appl Environ Microbiol*, 63(10):3985–3991, 1997.

[144] A. Molina-Höppner, T. Sato, C. Kato, M.G. Gänzle, and R.F. Vogel. Effects of pressure on cell morphology and cell division of lactic acid bacteria. *Extremophiles*, 7:511–516, 2003.

[145] D. Knorr, V. Heinz, and R. Buckow. High pressure application for food biopolymers. *Biochim Biophys Acta*, 1764(3):619–631, 2006.

[146] F. Skouri-Panet, S. Quevillon-Cheruel, M. Michiel, A. Tardieu, and S. Finet. sHSPs under temperature and pressure: The opposite behaviour of lens alpha-crystallins and yeast HSP26. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, 1764(3):372–383, 2006.

[147] A. P. Hammersley. ESRF internal report ERRF97HA02T, 1997.

[148] P. Boesecke. Reduction of two-dimensional small- and wide-angle X-ray scattering data. *Journal of Applied Crystallography*, 40(s1):s423–s427, 2007.

[149] D.I. Svergun. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst*, 25:495–503, 1992.

[150] R. Lange, J. Frank, J.L. Saldana, and C. Balny. Fourth derivative uv-spectroscopy of proteins under high pressure. i. factors affecting the fourth derivative spectrum of the aromatic amino acids. *Eur Biophys J*, 24(5):277–283, 1996.

[151] D. Svergun, C. Barberato, and M.H.J. Koch. Crysol - a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Cryst*, 28:768–773, 1995.

[152] M. Fukuda and S. Kunugi. Pressure dependence of thermolysin catalysis. *Eur J Biochem*, 142:565–570, 1984.

[153] R. J. Frankenberg, M. Andersson, and D. S. Clark. Effect of temperature and pressure on the proteolytic specificity of the recombinant 20S proteasome from *Methanococcus jannaschii*. *Extremophiles*, 7(5):353–360, 2003.

[154] M. Tehei, B. Franzetti, D. Madern, M. Ginzburg, B.Z. Ginzburg, M.-T. Giudici-Orticoni, M. Bruschi, and G. Zaccai. Adaptation to extreme environments: macromolecular dynamics in bacteria compared *in vivo* by neutron scattering. *EMBO Rep*, 5(1):66–70, 2004.

Chapter A

Annexes

A-1 SANS calculations

Intensities at q=0 of sample 1 and sample 3 (see section 4.3.2), that are extrapolated from the Guinier plot assuming that there are no interactions between the protein complexes in solution, are:

$$\text{Sample 1: } I_{0,1} = 1.01 \quad \text{Sample 3: } I_{0,3} = 0.684$$

Furthermore the protein concentration c_0 in mg/ml is the same in both samples and for the sake of simplicity is assumed to be unity. It is assumed that sample 1 contains only dodecamers and that only two species are present in sample 3. The unknown dimer concentration in sample 3 is denoted c_2 and the unknown dodecamer concentration c_{12}. Therefore:

$$1 = c_0 = c_2 + c_{12} \quad \Rightarrow \quad c_2 = 1 - c_{12}$$

The molecular mass of the dodecamer is designated M_{12} and therefore the molecular mass of the dimer is $\frac{1}{6}M_{12}$. It has been shown by Jacrot and Zaccai [119] that $I_{0,1}$ and $I_{0,3}$ can then be expressed as:

$$\begin{aligned} I_{0,1} &= A_1 \left(M_{12}c_0\right) \\ I_{0,3} &= A_3 \left(M_{12}c_{12} + \frac{1}{6}M_{12}c_2\right) \end{aligned}$$

The proportionality constant A_1 and A_3 are the same and so the dimer concentration c_2 can be calculated by dividing $I_{0,3}$ by $I_{0,1}$:

$$\begin{aligned}
\frac{I_{0,3}}{I_{0,1}} &= \frac{c_{12} + \frac{1}{6}c_2}{c_0} \\
&= 1 - c_2 + \frac{1}{6}c_2 \\
\frac{5}{6}c_2 &= 1 - 0.677 \\
c_2 &= 0.39
\end{aligned}$$